Université de Marne-La-Vallée

Institut Géographique National
Laboratoire COGIT

THÈSE

pour obtenir le grade de

Docteur de l'Université de Marne-La-Vallée

Spécialité : Sciences de l'Information Géographique

présentée et soutenue publiquement par

Patricia Bordin

le 12 décembre 2006

Méthode d'observation multi-niveaux pour le suivi de phénomènes géographiques avec un SIG

A Multi-Level Observation Method for a Follow-up Report of Geographical Phenomena in a GIS

Directeur de thèse
Léna Sanders

Jury

Saint-Julien Thérèse	rapporteur
Voiron Christine	rapporteur
Sanders Léna	directeur de thèse
Deutsch Jean-Claude	examinateur
Le Men Hervé	examinateur
Salgé François	examinateur

© UMLV

Remerciements

Je remercie le laboratoire COGIT et son chef, Anne Ruas de m'avoir accueillie pendant ces recherches. Je remercie plus globalement l'IGN, en particulier Jean Denègre alors directeur de l'ENSG, de m'avoir poussée dans l'expérience d'une thèse, Jacques Poulain et la direction technique pour la confiance accordée pour mener ces recherches. Cette expérience me permet de proposer des éléments de réponse aux thématiciens ; ils ont indirectement contribué à l'élaboration de cette thèse en me confrontant à leurs problématiques lorsque j'étais responsable de l'enseignement en SIG à l'ENSG.

Je remercie Léna Sanders, ma directrice de thèse qui a toute mon estime et ma reconnaissance. Elle m'a accompagnée dans mes recherches et m'a aidée à aller plus loin. Ses relectures précises et attentives, ses critiques toujours enrichissantes ont permis une formalisation plus rigoureuse.

Je remercie Jean-Claude Deutsch, qui le premier a vu l'intérêt d'une recherche à *l'entre-deux* et m'a donné la chance de tester et de valider mes travaux. Je le remercie pour son amitié et pour le temps passé à échanger autour de la thèse ; ces moments m'ont aidée tout au long de mon cheminement.

Je remercie mes deux rapporteurs : Christine Voiron, qui m'a proposé de faire des liens avec les outils d'analyses morphologiques. J'aimerais le faire avec elle comme suite à la discussion entamée lors de la soutenance. Je remercie Thérèse Saint-Julien, son rapport est pour moi une grande source d'encouragements. Le profond respect que je lui porte rend ce soutien plus précieux encore.

Je remercie les autres membres de mon jury : François Salgé de me conforter dans mes convictions, son parcours professionnel entre IGN et utilisateurs donne à sa parole sur cette thèse une importance particulière ; Hervé Lemen d'avoir été aux origines de cette thèse et d'avoir accepté de participer à sa présentation finale. Il m'importait d'avoir son regard sur ces travaux.

Je remercie Olivier Bonin pour son aide pour préparer l'oral et chaque fois que j'ai fait appel à lui. Je remercie Nadine Cattan, Renaud le Goix et Bruno Tassin, ils ont beaucoup contribué à ma soutenance par leurs remarques et leurs conseils, mais aussi par l'amitié avec laquelle ils me les ont donnés.

Je remercie François Durand Dastès qui m'a aidée à éclaircir certaines notions. Notre première discussion est un des moments charnières de la thèse.

Je remercie Philippe Fontaine qui par ses conseils m'a permis d'adapter la forme pour qu'elle ne devienne pas un poids pour le fond.

Je remercie Maria Basile qui m'a offert, avec l'ambiance urbaine, un objet géographique pour valoriser la méthode d'observation proposée. Je remercie Marie Mayeux et Olivier Raymond qui ont contribué à cette thèse par la qualité de leur travail de stage sur le coefficient d'imperméabilisation et les zones pavillonnaires. Je remercie Hélène Mathian pour le travail en commun sur les zones pavillonnaires ; les premiers résultats ont servi à initier l'application principale de ces recherches. Je remercie Arnaud Braun, Cécile Duchène de m'avoir fait bénéficier de leurs travaux en me fournissant des données utiles à l'application de ces recherches, Sébastien Mustière pour son expertise en multi-échelleS qu'il a bien voulu me faire partager. Je remercie également Jean-François Gleyze sans qui ces années de recherche auraient été nettement plus difficiles, Benoit Poupeau, ainsi que mes autres collègues du COGIT.

Je remercie pour leur soutien administratif et logistique Mounia Ahmedi, Marie-Claude Foubert et Alain Sombris dont la gentillesse légendaire n'est pas qu'une simple réputation.

Je remercie ceux qui m'ont aidée dans ces recherches à divers titres : géographes, informaticiens, géomaticiens, thématiciens, chercheurs ou opérationnels,… je ne peux tous les citer, pour ne pas omettre de nom, je les remercie globalement avant de pouvoir le faire lors d'une prochaine rencontre.

Enfin, je remercie Valérie Barrière ma fidèle lectrice, je ne saurais lui dire suffisamment merci.

Merci à toute ma famille pour sa tendresse et son indéfectible soutien,
à mes parents, à ma sœur d'avoir été et d'être toujours là,
à Noëlle et Yolande, pour être là maintenant,
à Paul, Tom et Duncan qui n'aspiraient qu'à l'*après-thèse*, mais ont fait tout ce qui était en leur possible pour me faciliter le *pendant la thèse*. Je leur dis ici combien je les aime.
La soutenance de la thèse leur était dédiée…

Enfin, à Bruno,… tout simplement pour tout.

Résumé

Le suivi de phénomènes géographiques avec un système d'information géographique (SIG) confronte l'utilisateur au problème de la gestion de la dimension temporelle. Des solutions existent ; la plupart sont techniques et spécifiques. Elles requièrent généralement des développements informatiques à façon, fonction des phénomènes considérés.

Cette thèse développe une solution générique, formelle, fondée sur une méthode d'observation multi-niveaux. Elle travaille pour cela sur une étape particulière du processus de modélisation des *phénomènes géographiques* à l'aide de *données géographiques* ; cette étape se situe entre la formalisation thématique et la formalisation technique.

La méthode d'observation multi-niveaux propose de concevoir des indicateurs détaillés, simples à suivre dans un SIG et compatibles avec des informations existantes. Elle s'appuie sur une méthode d'agrégation originale qui lui permet d'intégrer des données de sources multiples pour créer des indicateurs relevant de niveaux différents et de domaines de caractérisation variés.

L'objectif de la méthode d'observation multi-niveaux est de permettre aux utilisateurs de profiter des fonctionnalités d'analyses des SIG pour construire des informations utiles au suivi des phénomènes qu'ils étudient, à leur analyse, voire dans une démarche plus prospective à leur anticipation. La méthode est une proposition pour des outils SIG de type tableau de bord et outils de veille, qui pourraient ainsi aider dans les décisions.

Mots clés : SIG, observation mutli-niveaux, suivi d'évolutions, phénomènes géographiques, aide à la décision.

Abstract

The follow-up report of geographical phenomena by use of a geographical information system (GIS) confronts the user with the problem of time management. There are ways to overcome this. Most of them are technical and specific. On the whole, these require custom-made computer science developments.

This doctoral thesis develops a formal generic solution based on a multi-level observation method. To that end it focuses on a particular stage of the modelling process of geographical phenomena using geographical data; this stage occurs between the thematic formalization and technical formalization.

The multi-level observation method suggests devising detailed indicators simple to follow in a GIS and compatible with existing information. It is based on an original aggregation method which allows to integrate data from many sources in order to create indicators relating to different levels and various characterization fields.

The object of the multi-level observation method is to allow the user to benefit from GIS function analyses to build up useful information for a follow-up report of phenomena he/she is studying, for their analysis, indeed, to anticipate them in a more prospective approach thus supporting them with their decision-making.

Keywords: GIS, multi-level observation, development follow-up report, geographical phenomena, decision-making support

Sommaire

Introduction 15 - 20

Partie A :

Eléments de problématique

A.1. Les SIG : des outils dédiés aux informations géographiques 21 - 40

A.1.1. Définitions préalables 21

A.1.2. SIG et informations géographiques 27
A.1.2.1. Les SIG comme SI pour les informations géographiques 27
A.1.2.2. Deux besoins à l'origine des SIG, le SIG au croisement de deux cultures 28
A.1.2.3. L'approche par champ et l'approche par objet 28
A.1.2.4. Les composantes de l'information géographique 29
A.1.2.5. La modélisation des informations géographiques en base de données 33

A.1.3. De l'analyse dans les SIG aux SIG outils d'aide à la décision 35
A.1.3.1. Les utilisations des SIG 35
A.1.3.2. Les SIG inventaires/observatoires comme préalable aux analyses dans les SIG 36
A.1.3.3. Vers un SIG, outil d'aide à la décision 38

A.1.4. Intégration de la dimension temporelle dans l'observation des informations géographiques 39

A.2. Études de phénomènes géographiques 41 - 66

A.2.1. Observation de phénomènes 41
A.2.1.1. Phénomènes : éléments de définition 41
A.2.1.2. L'approche continue, l'approche discrète 41
A.2.1.3. Représentation discrète du mouvement, suivi d'évolutions à l'aide de mutations 44
A.2.1.4. Observation de la dimension spatiale et de la dimension temporelle 47

A.2.2. Etude de phénomènes 52
A.2.2.1. Etudes de la dimension spatiale 52
A.2.2.2. Etude de la dimension temporelle : les phénomènes évolutifs 56
A.2.2.3. Etudes des évolutions 58
A.2.2.4. Les phénomènes dits géographiques 61

A.2.3. Vers une méthode d'observation des phénomènes géographiques à l'aide de SIG 65

A.3. Observation temporelle dans les SIG 67 - 90

A.3.1. Observations temporelles et SIG 67
A.3.1.1. Observation de la dimension temporelle 67
A.3.1.2. Observation de phénomènes géographiques 70

A.3.2. Intégration de la dimension temporelle dans un SIG 70
A.3.2.1. Le temps dans les SIG : point de vue conceptuel 70
A.3.2.2. La mise à jour : point de vue technique 71
A.3.2.3. La gestion des mises à jour : point de vue pratique 72

A.3.3. Vers une gestion élaborée des mises à jour dans un SIG 86
A.3.3.1. Approche technique : un exemple de configuration pour gérer la dimension temporelle 87

A.3.3.2. Approche méthodologique : vers une méthode pour élaborer des données de suivi dans un SIG 88
A.3.3.3. Trois contextes applicatifs différents pour tester la méthode proposée 89

Partie B :

Eléments méthodologiques

B.1. Entre données et phénomène observé : la formalisation des informations géographiques 91 - 114

B.1.1. L'information géographique comme médiateur entre une approche thématique et une approche technique 91
B.1.1.1. Entre phénomène et données : l'information géographique 91
B.1.1.2. Deux informations géographiques : celle du technicien, celle du thématicien 93
B.1.1.3. L'*entre-deux*, entre technique et thématique 95

B.1.2. A l'entre-deux : la formalisation technique 100
B.1.2.1. L'étape de formalisation technique 101
B.1.2.2. Différenciation des informations géographiques 104
B.1.2.3. Informations à observer, informations observables : rôle de la modélisation en 'objet informatique' illustré sur deux exemples 108

B.1.3. Vers l'observation de phénomènes géographiques avec un SIG 114

B.2. Méthode d'observation multi-niveaux dans un SIG et méthode d'agrégation par emprise 115 - 168

B.2.1. Observation statique de phénomènes géographiques dans un SIG 115
B.2.1.1. Observation d'*objets géographiques* 116
B.2.1.2. Définition de l'objet d'observation principal et niveau d'observation 116

B.2.2. Vers une observation multi-niveaux dans un SIG 120
B.2.2.1. La gestion de multiples représentations dans les SIG 120
B.2.2.2. Solutions pour gérer des représentations multiples 122

B.2.3. Vers une méthode de création d'informations multi-niveaux pour l'observation de phénomènes 125

B.2.4. Méthode d'agrégation par emprise dans un SIG 128
B.2.4.1. Les méthodes d'agrégation 128
B.2.4.2. Caractérisation des méthodes d'agrégation d'objets élémentaires 129
B.2.4.3. Méthode d'agrégation par emprise 130
B.2.4.4. La méthode d'agrégation par emprise pour une méthode d'observation multi-niveaux dans les SIG 133

B.2.5. Méthode d'observation multi-niveaux dans un SIG 134
B.2.5.1. Principe de la méthode d'observation multi-niveaux 134
B.2.5.2. Cas où l'emprise d'agrégation n'est pas donnée 136
B.2.5.3. Caractérisation des 'agrégations d'objets micro' 148

B.2.6. Extensions de la méthode d'observation multi-niveaux 155
B.2.6.1. Enrichissement de la composante sémantique des 'objets' utilisés 156
B.2.6.2. Exploitation d'analyses morphologiques 158
B.2.6.3. Exploitation d'analyses topologiques 160
B.2.6.4. Exploitation d'analyses statistiques 162
B.2.6.5. Utilisation de plusieurs agrégations 164

B.2.7. De l''objet méso-structurel' à la 'portion de territoire' 165

B.3. Suivi de phénomènes géographiques avec un SIG 169 - 198

B.3.1. Suivi d'évolutions de phénomènes géographiques 169
B.3.1.1. Evolution et changement 170
B.3.1.2. Suivi de phénomènes géographiques dans un SIG 173

B.3.2. Suivi de phénomènes géographiques dans un SIG avec la méthode d'observation multi-niveaux 176
B.3.2.1. Exploitation standard de la méthode 176
B.3.2.2. Amélioration de la méthode d'observation multi-niveaux 187

Partie C :

Eléments d'application

C. Cas de mise en œuvre et d'application 199 - 234

C.1. Cas de phénomènes géographiques 199

C.2. Le phénomène d'urbanisation pavillonnaire 202
C.2.1. Observations sur le phénomène d'urbanisation pavillonnaire 205
C.2.2. Mise en œuvre de la méthode pour l'observation des zones pavillonnaires 208
C.2.3. Vers un outil d'observation et d'analyse de l'urbanisation pavillonnaire 212

C.3. Le phénomène d'imperméabilisation des sols 214
C.3.1. Suivi de l'imperméabilisation des sols pour réduire les risques liés au ruissellement 214
C.3.2. Vers un suivi du coefficient d'imperméabilisation 216
C.3.3. Mise en œuvre de l'observation de l'imperméabilisation des sols 217
C.3.4. Vers un outil opérationnel d'observation de l'imperméabilisation des sols au service des aménageurs 219

C.4. Le phénomène d'ambiance urbaine 221
C.4.1. Définition de la notion d'ambiance, différentes approches 221
C.4.2. Vers une observation globale de l'ambiance urbaine 223
C.4.3. Mise en œuvre de l'observation de l'ambiance urbaine 227
C.4.4. Vers un outil opérationnel d'observation de l'ambiance urbaine au service des collectivités 229

C.5. Conclusion sur les trois déclinaisons de la méthode d'observation générique 232

Conclusion et perspectives 237 - 242

Bibliographie 243 - 250

Annexes 251 – 283

Annexe 1 : Résolution, précision et granularité : présentation comparée 252

Annexe 2 : Les *thèmes* : une approche technique pour spécifier les informations géographiques 254

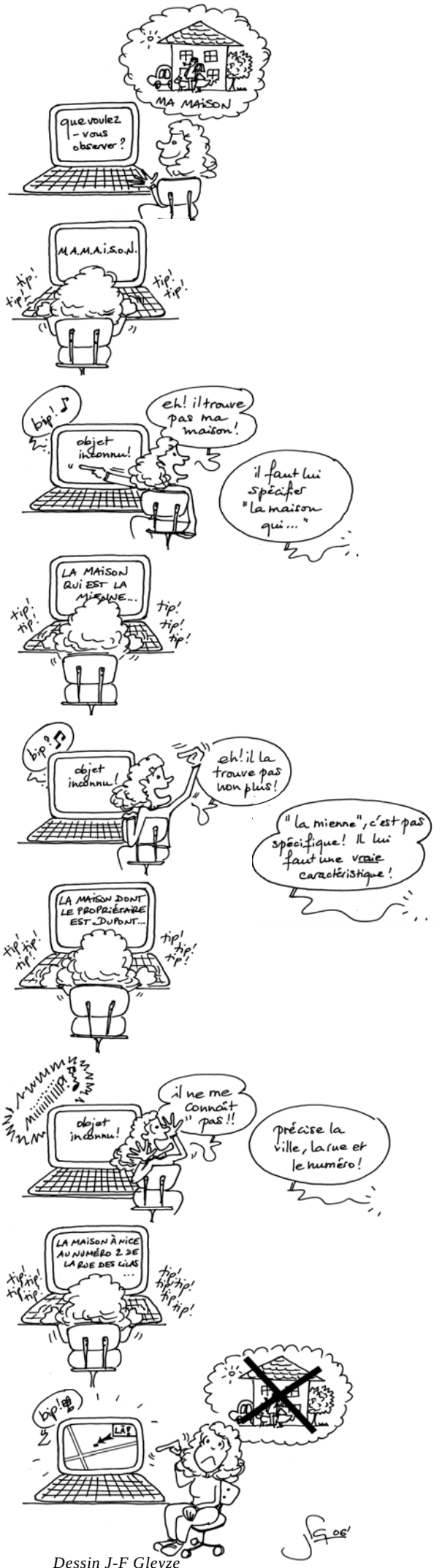
Annexe 3 : Définition de l'*entre-deux* 257

Annexe 4 : Une représentation des 'objets' sous forme d'information par nature pour une observation simplifiée dans les SIG 258

Annexe 5 : Eléments de discussion sur l'hypothèse d''emprise au sol constante 260

Annexe 6 : Formulation mathématique des indicateurs produits avec la méthode d'observation multi-niveaux 262

Annexe 7 : Eléments d'applications sur les cas de phénomènes retenus 265
Annexe 8 : Etapes de mise en œuvre d'une méthode d'observation multi-niveaux 279


MA MAISON
que voulez-vous observer ?
M.A.M.A.i.S.O.N.
tip! tip! tip!
bip!
objet inconnu!
eh! il trouve pas ma maison!
il faut lui spécifier "la maison qui..."
LA MAISON QUI EST LA MIENNE...
"la mienne", c'est pas spécifique! Il lui faut une vraie caractéristique!
LA MAISON DONT LE PROPRIÉTAIRE EST DUPONT...
Mmmmmm?
il ne me connaît pas!!
précise la ville, la rue et le numéro!
LA MAISON À NICE AU NUMÉRO 2 DE LA RUE DES LILAS ...
LAT

Dessin J-F Gleyze
Scénario P.Bordin 2006

Introduction

I.1. Observation de phénomènes géographiques avec un SIG

Les SIG des outils prometteurs

Depuis leur origine dans les années 70, les SIG (système d'information géographique) sont accompagnés de la promesse d'analyses géographiques sophistiquées. Mais plusieurs décennies plus tard, force est de constater que ces promesses, toujours actuelles, ne se sont pas vraiment transformées dans la pratique des thématiciens : les SIG sont souvent cantonnés à leur rôle originel d'outils à la fois de gestion et de représentation des informations géographiques. Ce rôle, sur lequel ils ont la quasi-exclusivité, leur permet de produire des résultats essentiels pour les thématiciens sur les phénomènes (localisation, concentration, dispersion, spécialisation). Ces résultats semblent suffire à justifier leur existence. Pourtant les fonctionnalités qu'ils offrent en standard pour analyser des aspects géométriques et topologiques et surtout pour les analyser conjointement avec des aspects descriptifs, leur permettraient d'être plus que de simples « meubles à plans sophistiqués ». Dans la lignée des outils décisionnels, ils pourraient aider aux décisions portant sur des aspects géographiques. Toutefois pour cela, il faudrait déjà qu'ils aient satisfait les besoins thématiques plus essentiels : avant de pouvoir décider, l'utilisateur à en effet besoin de connaître, de comprendre et pour cela d'observer.

Un besoin d'aide à l'observation

Le discours des thématiciens révèle en effet un besoin de *moyens*, de *méthodes* pour étudier les phénomènes qui les intéressent *« il s'agit déjà de trouver le moyen de lire les processus, de comprendre les mécanismes sous-jacents à certaines situations, d'établir une sorte de protocole d'approche temporelle » {Basile, 2005}.*

« Dans ce monde qui nous submerge d'informations, on comprend aisément pourquoi il est possible de se sentir perdre pied à l'intérieur de son propre domaine... D'où le besoin accru de structures de références

susceptibles de faciliter l'intégration de la masse d'informations toujours changeantes que l'homme doit traiter » [Hall, 1971].

Pour les informations géographiques, ne serait-ce pas justement là la vocation des SIG ? La localisation y sert de *référence* pour intégrer une masse d'informations (par géocodage), qui sont *structurées* en représentations à des échelles différentes. Ils offrent des fonctionnalités de *traitement* et leur capacité de stockage est *quasi-illimitée* (du fait de leur lien avec les systèmes de gestion de bases de données). En revanche, ils peinent effectivement à intégrer le fait que les informations soient « toujours changeante » …

Pourquoi alors ce besoin de moyens d'observation alors que les SIG semblent là pour y répondre ? Deux raisons peuvent l'expliquer :

- les SIG restent des outils informatiques conçus par des spécialistes, ce qui restreint leur accessibilité pour les utilisateurs,
- l'observation est une application complexe à mettre en œuvre car elle est elle-même évolutive.

Les SIG des outils informatiques spécialisés

Les SIG malgré leur habillage ergonomique relativement récent, restent des outils fondamentalement complexes. La modélisation, la production, la gestion des données géographiques nécessitent des compétences spécifiques qui ne sont pas contenues dans le mode d'emploi des logiciels de SIG. L'implémentation des données dépend de plus du phénomène observé, du point de vue adopté ; elle demande souvent un travail à façon pour adapter les moyens techniques aux besoins de l'utilisateur. Plus encore, certains aspects étudiés par les thématiciens ne peuvent être restitués dans les outils SIG qu'au prix de solutions techniques complexes, difficiles à mettre en œuvre par les utilisateurs ; c'est en particulier le cas de la dimension temporelle.

L'observation est une application complexe

On peut distinguer plusieurs types d'observations : la plus simple, l'observation statique, initie le processus d'acquisition de connaissances, mais celui-ci se perfectionnant le besoin d'observation devient alors de plus en plus dynamique et donc de plus en plus complexe. Le processus d'acquisition de connaissance étant généralement continu, le besoin de l'utilisateur porte donc sur une solution évolutive capable de l'accompagner dans ses avancées.

L'observation une application stratégique pour les utilisateurs et donc pour les SIG

Or les SIG ne peuvent pas aujourd'hui apporter de solutions simples pour satisfaire ce besoin d'évoluer vers des analyses dynamiques. En effet, s'ils répondent parfaitement aux besoins d'observations statiques de type : inventaire d'objets géographiques présents sur un territoire, la modélisation et les fonctionnalités standards des SIG limitent en revanche leurs capacités à effectuer des analyses dynamiques pour observer les processus en jeu dans un phénomène.

De fait, l'observation temporelle dans les SIG se réduit souvent à l'affichage simultané de deux images (cartes ou photographies aériennes) et à une simple comparaison visuelle. Des solutions existent certes pour identifier automatiquement les évolutions, mais ces solutions sont informatiques et souvent difficiles à mettre en œuvre par un non-informaticien. Par ailleurs, elles délèguent l'interprétation des évolutions à l'ordinateur. Or cette interprétation dépend du contexte (ex. : la création d'un bâtiment n'a pas le même sens dans une zone urbanisée et dans une

zone d'espaces verts protégés), du point de vue (la création d'un bâtiment prend un sens différent en urbanisme, en sociologie, en hydrologie, etc.) et même de la sensibilité de l'observateur (ex. : combien de création de bâtiments est nécessaire pour qu'un espace agricole soit considéré comme espace urbanisé ?). Ces solutions ne permettent pas d'observer des nuances, de saisir les subtilités utiles à la prise de décisions.

Une autre application des SIG de type observation est toutefois envisageable, qui pourrait également accompagner l'utilisateur dans le processus d'aide à la décision. Elle ne nécessite pas de développer les capacités d'*analyses dynamiques* des SIG, mais seulement de développer leurs capacités à restituer aux utilisateurs les informations dont ils ont besoin pour les réaliser. Elle délègue à l'utilisateur l'identification et l'analyse dynamique des processus, mais elle lui fournit plus d'informations et des informations adaptées pour le faire. Cette application qui ressemble à celle d'un tableau de bord est essentielle aux outils d'aide à la décision ; dans les outils décisionnels, cette fonction est d'ailleurs assurée par l'EIS (*executive information system*). Permettre aux SIG d'assurer cette fonction, les aiderait à évoluer du statut de structure informatique de gestion des informations géographique vers le rôle plus stratégique d'outils de suivi de phénomènes géographiques. Par ailleurs, en développant leur capacité à fournir à la fois des informations plus nombreuses et plus détaillées, sous forme d'indicateurs simples à suivre dans le temps, ils pourraient aider les thématiciens à prendre de meilleures décisions (car fondée sur une connaissance plus détaillée), à les prendre au bon moment (grâce à un suivi « en temps réel »), et à les suivre dans le cadre d'un processus de feed-back pour les corriger. Le SIG serait conçu comme un outil de veille sur le territoire : il fournirait des indicateurs pour suivre les évolutions, pour mesurer les écarts entre les résultats attendus et ceux obtenus, pour observer les éventuels effets secondaires de ces décisions, et en régime stationnaire, pour voir apparaître des informations aberrantes comme d'éventuels signes précurseurs de phénomènes émergents.

Cette fonction de veille ne se confond pas avec celles de gestion et de restitution des informations géographiques (dans les outils décisionnels ces fonctions de gestion sont d'ailleurs assurées par un autre système que l'EIS : le datawarehouse). La fonction de veille implique de mettre en œuvre des fonctionnalités d'analyses (morphologiques, topologiques, statistiques) pour construire les indicateurs. Mais, surtout elle requiert des capacités de gestion temporelle plus élaborées que celles disponibles en standards dans les logiciels de SIG.

Objectifs fixés pour la thèse

Dans cette thèse nous proposons de construire une méthode générique pour permettre aux utilisateurs de décliner des méthodes spécifiques de veille, pour suivre les phénomènes particuliers qu'ils étudient.

Nous partons pour cela d'une hypothèse : la technicité[1] des SIG les tient trop éloignés des thématiciens. Pour la plupart des utilisateurs, le chemin qui les conduirait vers les solutions informatiques pour le suivi des phénomènes dans un SIG est en dehors de leurs domaines de compétences centrés autour des phénomènes qu'ils analysent. Il en résulte généralement des difficultés à mettre en œuvre un suivi des phénomènes adapté ; les observations sur les évolutions sont alors souvent trop sommaires pour alimenter efficacement les modèles d'analyses thématiques développés par ailleurs par les utilisateurs. Nous stipulons que ce ne doit pas être aux utilisateurs d'aller vers les solutions techniques ; c'est aux SIG de venir aux utilisateurs et de les conduire petit à petit vers des solutions de plus en plus

1. Technique étant ici à entendre au sens de *« qui appartient à un domaine particulier, spécialisé, de l'activité ou de la connaissance »*.
« Cette origine « technicienne » a profondément marqué le développement des technologies de l'information géographique, non seulement en termes de domaines d'utilisation mais également dans l'ergonomie et dans la conception même des logiciels et des modèles de données » [Feyt, 2004]

techniques certes, mais également de plus en plus performantes. Nous proposons pour cela une solution alternative au problème de suivi avec un SIG : une solution méthodologique et formelle. Cette thèse élabore ainsi une méthode qui permette aux utilisateurs de s'affranchir de façon générique des contraintes posées par les SIG pour observer des phénomènes géographiques.

Le caractère générique de la méthode n'est pas la seule condition que nous imposons a priori à ces recherches. En effet :

– nous voulons également que les méthodes spécifiques déduites de la méthode générique aient la capacité d'accompagner les thématiciens dans le double processus d'acquisition de connaissances :

- acquisition de connaissance sur les phénomènes qu'ils étudient
- et sur les outils (SIG) qu'ils utilisent.

Pour répondre à ce besoin nous avons cherché à construire une méthode *adaptative*, c'est à dire une méthode qui s'adapte à des utilisateurs ayant des niveaux de compétences-SIG différents. Cette qualité en retour assure aux méthodes spécifiques de pouvoir être *évolutives* dans la durée de l'application et donc de pouvoir effectivement accompagner l'utilisateur dans ses avancés. En effet, si la méthode générique peut s'adapter à des niveaux de compétences différents, elles peut alors à l'inverse utiliser cette capacité pour proposer aux thématiciens d'intégrer des modèles et des outils plus élaborés pour faire évoluer leur méthode d'observation ; ils disposeront alors de moyens supplémentaires pour acquérir de nouvelles connaissances sur les phénomènes qu'ils analysent.

– plus généralement, la méthode générique d'observation doit être :

- ouverte sur une large gamme de phénomènes, ouverte sur des profils d'utilisateurs différents, mais aussi,
- ouverte sur l'existant, en particulier, les méthodes d'observation spécifiques devront prendre en compte les applications existantes et se placer en complémentarité,
- ouverte sur les modèles de gestion temporelle, pour intégrer les informations d'évolutions qu'ils fournissent,
- ouvertes sur les méthodes et outils d'analyses d'informations géographiques qui peuvent exister en dehors des fonctionnalités standards des SIG, pour profiter de leurs capacités à créer des indicateurs supplémentaires,
- ouverte sur les méthodes d'analyses thématiques de phénomènes, afin de fournir des indicateurs proches de ceux utilisés par les thématiciens. En particulier, nous nous sommes inspirés des méthodes d'analyses multi-niveaux en géographie pour développer la méthode d'observation multi-niveaux proposée dans cette thèse.

Ces contraintes viennent ainsi s'ajouter à la contrainte principale posée par le sujet de cette thèse : le *suivi* de phénomène *dans le temps avec un SIG*.

I.2. Vers une méthode d'observation multi-niveaux pour suivre des phénomènes géographiques dans un SIG

Cette thèse qui vise à l'élaboration d'une méthode d'observation multi-niveaux pour les SIG se compose de trois parties.

La partie A pose la problématique, en rapprochant l'utilisation des systèmes d'information géographiques et l'étude de phénomènes géographiques (respectivement chapitre A.1 et A .2). Le chapitre A.3 présente les solutions techniques existant dans les SIG pour intégrer la dimension temporelle nécessaire à l'observation des phénomènes. Il montre en particulier les différents modèles disponibles, du plus simple qui sous-traite entièrement la gestion du suivi à l'observateur, aux solutions les plus élaborées demandant des développements à façon. Il introduit l'idée d'une méthode d'observation conceptuelle qui exploiterait le modèle le plus élémentaire, comme une alternative aux solutions existantes. Revenant sur les modèles plus sophistiqués en fin de thèse, il y sera montré comment les intégrer pour améliorer encore les observations.

La partie B s'organise en trois chapitres, qui correspondent aux trois étapes de l'élaboration de la méthode d'observation multi-niveaux.

Le chapitre B.1 montre que plusieurs étapes de formalisation sont nécessaires pour permettre l'observation d'une entité géographique à partir de données géographiques dans un SIG. Parmi elles, la *formalisation technique* située entre la *formalisation thématique* et *l'implémentation informatique* joue un rôle charnière : elle articule deux approches différentes de l'information géographique. Considérant cette étape comme l'étape technique la plus proche des utilisateurs nous décidons d'y développer la méthode d'observation.

Le chapitre B.2 débute par un truisme : « avant de pourvoir effectuer des observations détaillées dans un SIG, il est nécessaire de disposer de données » et par le constat que souvent cette condition préalable n'est pas remplie. Avant de chercher une solution pour résoudre le problème de suivi temporel, il s'agit déjà de résoudre un autre problème lié au SIG : le manque de données adaptées au suivi des phénomènes étudiés. Le chapitre B.2 est ainsi consacré à l'élaboration d'une méthode d'observation statique permettant de construire des indicateurs utiles à un suivi circonstancié des phénomènes géographiques. Nous inspirant de pratiques en géographie nous proposons alors de développer une méthode d'observation statique multi-niveaux pour les SIG ; elle a pour objet la construction d'informations sur les entités à observer à partir d'entités décrites à un niveau plus fin. Proches des méthodes d'analyses thématiques et permettant de construire des informations détaillées, le principe d'une méthode multi-niveaux pour observer des phénomènes géographiques dans les SIG semble prometteur ; il est aussi complexe à mettre en œuvre, car il se heurte à son tour à une autre difficulté technique: la gestion de représentations multiples. Comme pour la gestion temporelle, il existe des solutions pour gérer de multiples représentations dans les SIG, mais leur simplicité est généralement inversement proportionnelle à leurs performances. Ainsi la deuxième partie du chapitre B.2 est-elle consacrée à l'élaboration d'une méthode d'agrégation (dite *méthode d'agrégation par emprise*), qui puisse participer à la mise en œuvre d'une méthode d'observation multi-niveaux dans un SIG.

Avec la méthode d'agrégation par emprise, l'utilisateur dispose d'une méthode d'observation multi-niveaux *statique* pour construire des indicateurs utiles à l'observation de phénomènes dans un SIG. Le début du chapitre B.3 présente l'intérêt pour les utilisateurs d'avoir des indicateurs détaillés qui décrivent des états et les risques inhérents à une pré-interprétation des évolutions. Dans un second temps, ce chapitre montre que les indicateurs produits avec la méthode d'observation proposée peuvent être facilement suivis dans un SIG, sous l'hypothèse que l'emprise au sol du phénomène est constante. Il explique également comment

améliorer le suivi en intégrant des observations statiques d'origines diverses.

La partie C enfin présente trois déclinaisons de la méthode d'observation générique proposée, correspondant au suivi de trois cas de phénomènes particuliers : l'urbanisation pavillonnaire, l'imperméabilisation des sols, le phénomène d'ambiance urbaine. Ces cas d'études retenus pour la diversité des situations qu'ils offrent (sujet, données disponibles, problématique sous-jacente, problèmes de modélisation induits et difficultés techniques à résoudre, profil des utilisateurs finaux, etc.) ont largement contribué à l'élaboration de ces recherches. Dans cette partie, ils nous servent à mettre en avant des points particuliers de la méthode de suivi et à illustrer la diversité des phénomènes qu'elle permet d'observer. Ils attestent également de l'intérêt général de la méthode au travers de l'intérêt montré par les thématiciens pour les mises en œuvre spécifiques dédiées au suivi de chacun de ces phénomènes.

La conclusion propose une synthèse des éléments posés dans cette thèse et présente certaines des perspectives qu'ouvrent ces recherches.

A.1. Les SIG : des outils dédiés aux informations géographiques

A.1.1. Définitions préalables

Les projets de SIG (ou systèmes d'information géographique) impliquent l'interaction de compétences variées techniques (en modélisation informatique, en programmation, en conception de réseaux, en administration de données, en intelligence artificielle, etc.) et thématiques (en urbanisme, en transport, en sociologie, en hydrologie, etc.). Cette pluridisciplinarité fait leur richesse mais aussi leur complexité. Ainsi, un même terme peut avoir des sens différents suivant le contexte. C'est d'ailleurs le cas pour le terme *SIG* lui-même (qui peut signifier un logiciel, une application, des données, etc.) [Maguire et al., 1991], [Nyman, 1998]. Face à de telles ambivalences :

– nous tenterons d'unifier les points de vue quand ils nous paraîtront identiques ou complémentaires, comme nous l'avons fait pour la définition des SIG dans [Bordin, 2002][1]. Ainsi dans ce qui suivra, nous définirons les SIG de façon très générale comme : « *des outils informatiques permettant d'effectuer des traitements divers sur des données à références spatiales* », en nous laissant la possibilité de préciser cette définition en fonction du contexte (par exemple lorsqu'il s'agira des logiciels).

– pour les autres termes équivoques, nous choisirons une de leurs définitions après l'avoir explicitée. En particulier, nous définissons de suite les expressions *objet géographique* et *information géographique*. Ces définitions sont en effet fondamentales pour cette thèse : elles dissipent le flou qui existe souvent entre entité / objet, objet / information, information / donnée. Elles aident à formaliser la méthode que nous avons développée et qui est articulée autour de la modélisation des *informations géographiques* pour suivre des *objets* (géographiques) à l'aide de *données* dans un SIG.

1. Dans le chapitre 4, *Vers une définition des SIG*, p. 93-110, nous montrons que les différentes définitions correspondent :
- soit à une *approche fonctionnelle* : la définition est fonction de l'utilisation, du « pour quoi faire ? »
- soit à une *approche descriptive* : la définition est fonction du contenu ou du « fait de quoi ? ». Pour cette dernière, nous avons identifié plusieurs variantes : le SIG comme base de données, comme chaîne de traitements, comme logiciel et matériel, et enfin comme organisation ou service.

objet géographique

L'usage du mot *objet* dans un projet SIG peut renvoyer suivant les cas à des notions très différentes : à un objet physique, à un objet conceptuel, à

un objet informatique... Nous appuyant sur la distinction faite par [Couclelis, 1994] entre *entités* et *objets* :

« Entities are things in the real world that can be perceived in our representation (our mental model). Entities become objects. Objects are representation of things in the real world »,

nous distinguons :

- les objets *réels* (les entités)
- des objets *conceptuels*[2]

Puis, nous inspirant de [Timpf, 1997] :

« Entities refers to the things in the real world whereas the term objects refers to things in the database »,

nous proposons d'introduire une troisième catégorie d'objets, en référence aux bases de données :

- les objets *informatiques*

La notion d'« objet » existe par ailleurs en informatique. Elle renvoie à deux types d'*objets* :

− soit à une classe (par exemple la classe des objets tronçons de route),

− soit à une instance de classe, c'est à dire à un objet particulier de cette classe, à une valeur spécifique de l'« objet générique » décrit par la classe. Par exemple, l'objet zone pavillonnaire « le verger » à Roissy en-Brie appartient à la classe *zone pavillonnaire*.

Le sens dépend du contexte et du type d'étude. Dans les études formelles et conceptuelles où l'on raisonne sur des schémas, l'objet renvoie généralement à une classe comme dans l'exemple du schéma UML[3] de la figure A.1.1a. Dans les études plus applicatives, l'objet renvoie généralement à une entité particulière, à une *instanciation*[4] (fig. A.1.1b).

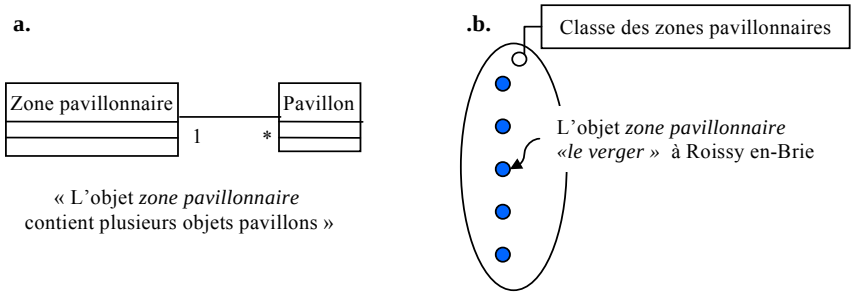


Figure A.1.1. *Deux sens pour la notion* d'objet *en informatique*

Dans cette thèse, sauf mention particulière, le terme « objet géographique» renverra à un *objet d'étude thématique,* c'est à dire à un objet tel qu'il est formalisé par un thématicien (par exemple un tronçon de route dans le cadre d'une étude de trafic, un bâtiment dans le cadre d'une étude architecturale, un espace vert dans le cadre d'un recensement agricole) et nous distinguerons trois types d'«objets » (fig. A.1.2):

− l'entité
− l'objet géographique
− l'objet informatique

- les **entités** désigneront des « objets réels », c'est à dire des objets appréhendés relativement au terrain physique.

- les **objets géographiques**, que nous noterons *objets géographiques* (en italiques), seront le résultat d'une sélection, d'une représentation et d'une conceptualisation par un thématicien.

- les **objets informatiques**, que nous noterons 'objets informatiques' (en utilisant un entre cote, en référence à la chaîne de caractères en informatique), seront le résultat de la formalisation d'*objets géographiques* pour les adapter à la modélisation propre aux SIG.

2. On peut discuter de l'existence de *l'objet réel* de façon philosophique. Toutefois, nos travaux se centrant sur la modélisation des informations géographiques et sur les données dans les SIG, ils n'abordent qu'indirectement les *entités*. Nous avons donc postulé que cette existence était vraie.

3. UML (unified modeling language) est un formalisme standard de modélisation, un langage graphique pour aider à l'analyse et à la conception orientées objet, défini par l'OMG (objet managment group) et issu des travaux de Grady Booch.

4. En informatique, *instancier* signifie donner une valeur particulière à un modèle générique ou encore, créer une donnée qui réponde au modèle.

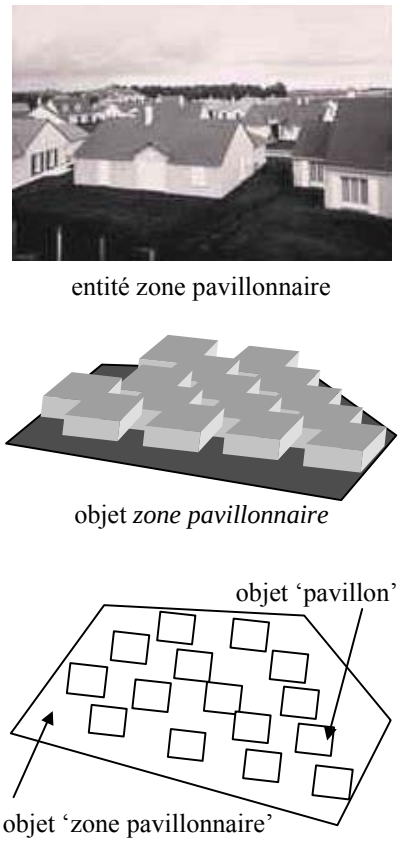


Figure A.1.2. *Zone pavillonnaire :* entité, *objet* et 'objet' *(photographie de Jürgen Nefzger)*

Pour ces travaux, il importe en effet de pouvoir distinguer la représentation d'une entité sous forme d'*objet géographique*, de la représentation sous forme d' 'objet informatique'.

EXEMPLE : une même entité géographique peut être appréhendée de façon différente.

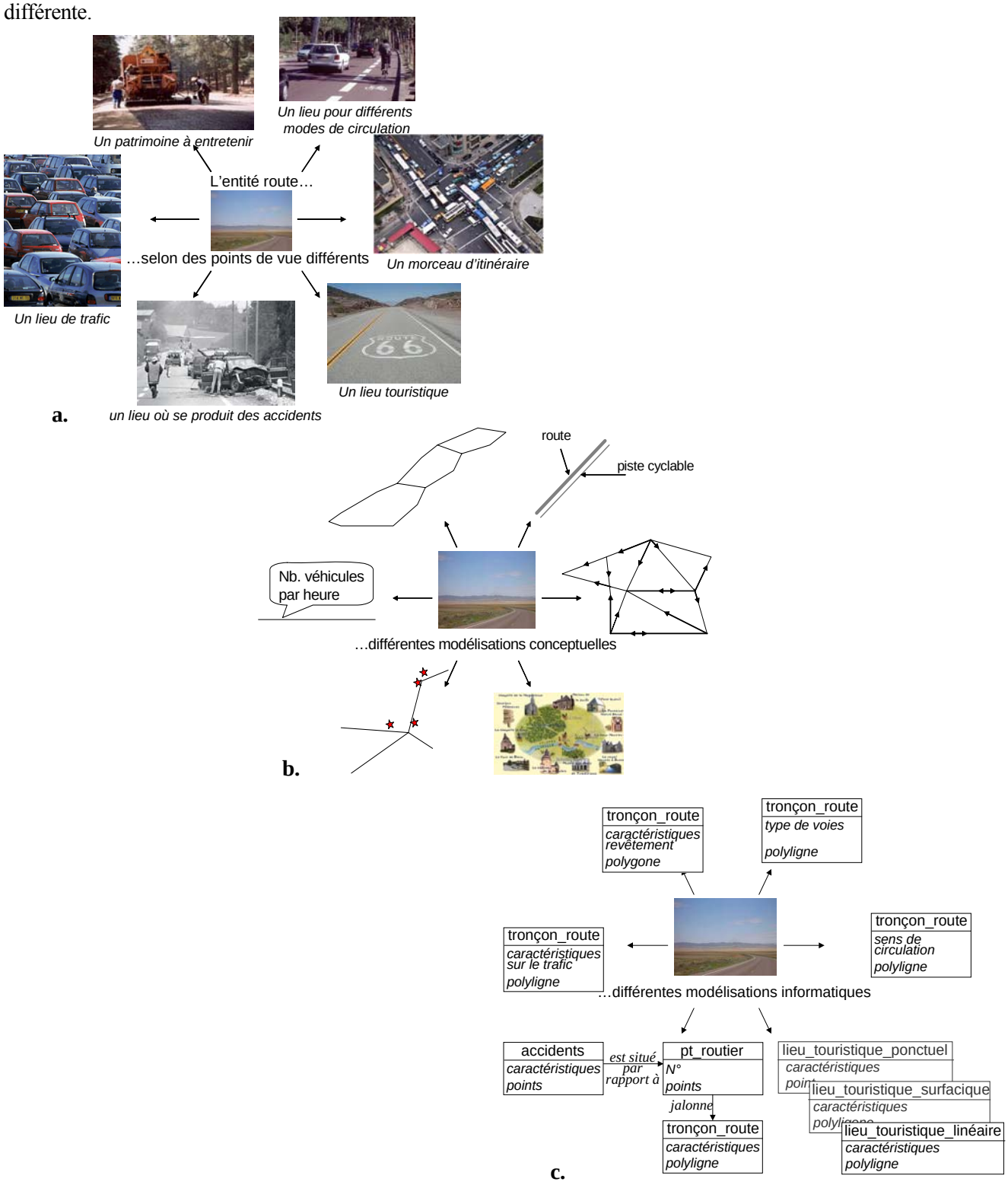


Figure A.1.3. *Comment une même entité physique peut être observée et représentée de façon différente*

Suivant le point de vue (ou les *catégories* au sens de système de pensée) du thématicien, les caractéristiques retenues ne sont pas les mêmes (fig. A.1.3a) (ex. : une *route* n'est pas considérée de la même façon dans une application sur la circulation et dans la gestion du revêtement). Il en résulte

généralement des formalisations différentes (fig. A1.3b) (ex.: l'objet géographique *route* sera envisagé comme un arc d'un graphe pour des études sur la circulation et comme une surface dans le cadre d'une application sur le revêtement). Le thématicien peut par ailleurs choisir d'utiliser un SIG (pour ses capacités de gestion des données, ainsi que pour ses capacités d'analyses) pour l'aider dans ses études. Il lui faut alors formaliser l'*objet géographique* en un 'objet informatique' pour l'adapter à la modélisation spécifique aux SIG (fig. A.1.3c). L'implémentation de ces 'objets informatiques' en données constitue encore une autre étape.

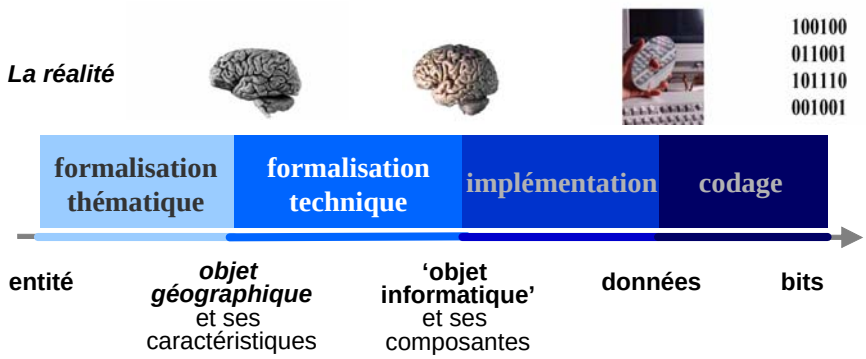


Figure A.1. 4. *De l'entité aux codes informatiques, différentes représentations*

Dans le processus de conceptualisation nécessaire à la mise œuvre d'une observation d'entités géographiques avec un SIG, nous proposons de distinguer plusieurs étapes (fig.A.1.4):

– la *formalisation thématique* part des entités ; elle sélectionne, caractérise, formalise des *objets géographiques*. Cette étape est celle qui permet au thématicien de modéliser les entités en {*objets (géographiques)* + caractéristiques} en fonction des observations et analyses qu'il veut réaliser.

– la *formalisation technique* consiste à modéliser des *objets géographiques* en 'objets informatiques' qui correspondent aux modèles utilisés dans les SIG. Cette étape s'appuie sur des compétences en *géomatique*[5] ; elle permet de définir des {'objets informatiques' + composantes} où les composantes comprennent des caractéristiques sémantiques d'une part et des caractéristiques sur la localisation des 'objets' d'autre part).

– l'*implémentation* adapte la modélisation des 'objets informatiques' aux particularités du modèle du logiciel-SIG retenu pour effectuer les traitements (ex. : codage explicite de la topologie dans ArcInfo, pas modélisation de la 3D dans MapInfo, types d'attributs gérés). Cette étape s'appuie sur des compétences en informatique et sur la connaissance du logiciel-SIG utilisé.

– après l'*instanciation* qui produit les données, le *codage* transforme les données entrées dans le logiciel en bits gérables par la machine. Cette étape revient à l'ordinateur.

Cette proposition de décomposition du processus intègre deux conceptions différentes :

– celle du producteur de données qui distingue et nomme :
- la *modélisation conceptuelle :* l'étape qui conduit de la réalité aux 'objets informatiques',
- la *modélisation logique* : l'étape d'implémentation en données dans un logiciel particulier,
- la *modélisation physique* : l'étape qui transforme les données en bits gérables par l'ordinateur.

5. La *Géomatique* est le nom donné au domaine regroupant les sciences et techniques opérant sur les informations géographiques, i.e. les méthodes et techniques de l'informatique appliquées aux moyens d'acquisition, de traitement, de représentation et d'archivage des données localisées.

– celle du thématicien qui différencie :
  - la *modélisation conceptuelle* : l'étape qui sélectionne, identifie, formalise et caractérise des *objets géographiques* à partir des entités,
  - de la *modélisation technique* : l'étape qui formalise ces *objets géographiques* pour aboutir à la création des données utiles aux traitements.

La figure A.1.5 qui rapproche ces deux points de vue, met en évidence une zone floue, qui semble n'être du ressort de personne (ou de tous) : celle que nous avons nommée : la *formalisation technique,* et qui fait la distinction entre *objet géographique* et 'objet informatique'. (Par exemple le thématicien qui observe une *route* a parfois du mal à envisager les conséquences d'une formalisation sous forme de surface par rapport à une formalisation sous forme de ligne, tandis que le technicien a du mal à déterminer quelles sont les informations importantes pour étudier un phénomène particulier). Ainsi, pour le technicien, la *formalisation technique* est incluse dans la modélisation conceptuelle ; pour le thématicien, elle fait partie de la modélisation technique. Or, comme nous le développerons dans le chapitre B.1, il existe bien une différence entre les deux objets et les distinguer permet de ne pas rigidifier la modélisation. Dans cette thèse nous proposons de profiter de cette souplesse pour élaborer plus d'informations à observer. Il nous faut pour cela définir auparavant le terme d' « information géographique».

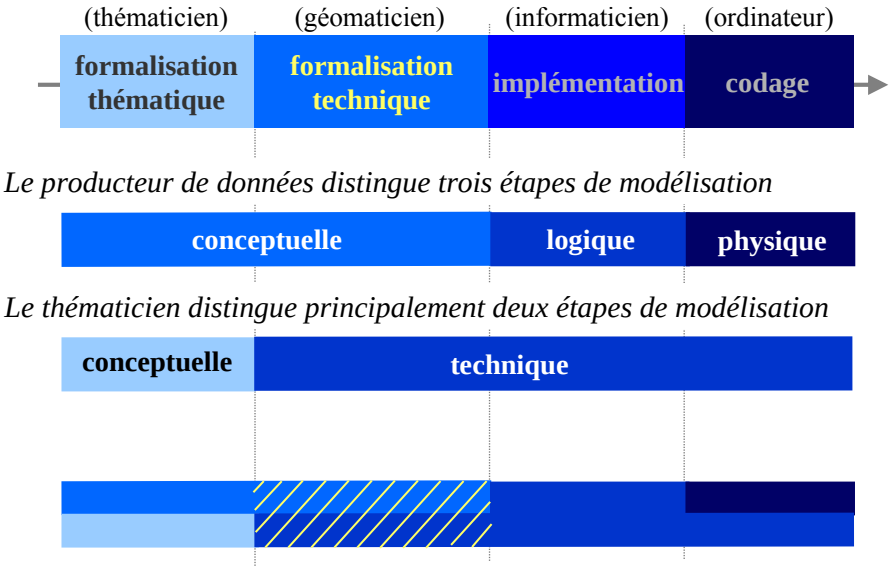


Figure A.1. 5. *L'étape de formalisation technique : une zone floue*

information géographique

Le terme d'information peut lui aussi prendre des sens différents. En particulier, on distingue :

- la définition du domaine informatique : *« faits et connaissances déduits des données. L'ordinateur manipule et gère des données. La signification* déduite *des données est l'information »* [Dictionnaire de l'informatique et d'internet]. Dans cette définition l'information est présentée comme provenant des données. Les données pré-existent aux informations.

- la définition générale du dictionnaire propose : *« élément, fait ou jugement sur la réalité que l'on peut transmettre »* [Petit Robert, 2001]. Cette fois, l'information semble se déduire de la réalité, les données ne sont plus évoquées explicitement et elles ne semblent pas devoir pré-exister aux informations. Dans le cadre de cette définition, on peut encore distinguer :

. l'information analytique et matérielle (ex . : « la maison est bleue »),
. du fait général, global, c'est à dire l'ensemble des apports sur un objet,
l'information comme un tout ou comme élément de connaissance.

Le fait de spécifier que l'information est «géographique» multiplie encore les définitions. En effet, outre les adaptations[6] des deux définitions générales précitées, on trouve diverses définitions particulières :

« Représentation d'un objet ou d'un phénomène réel, localisé dans l'espace à un moment donné » [Quodverte, 1994], cité par [Denègre et Salgé, 1996]

« L'information géographique peut-être définie comme une information relative à un objet géographique ou à un phénomène du monde terrestre, décrit plus ou moins complètement : par sa nature, son aspect, ses caractéristiques diverses, et par son positionnement sur la terre » (définition proposée par Denègre, sur le site du *serveur éducatif dédié à l'information géographique* : *www.seig.ensg.fr*).

« L'information géographique peut être définie comme l'ensemble de la description d'un objet et de sa position géographique à la surface de la Terre » (définition proposée par l'encyclopédie libre *http://fr.wikipedia.org)*.

« L'information géographique comprend toutes les données localisables sur le territoire, sous forme de points, de cartes, de tableaux, d'images, etc. » (définition du *Portail de l'information géographique gouvernemental* du Québec : *www.quebecgeographique.gouv.qc.ca)*

Certaines de ces définitions semblent se référer à des étapes différentes du processus de conceptualisation. Ainsi, la définition générale du dictionnaire et celle de [Quodverte, 1994] qui évoquent la réalité, pourraient intervenir dans l'étape de *formalisation conceptuelle* ; les « informations géographiques » seraient alors à des éléments (ou à la totalité) de connaissances spécifiés sur l'*objet géographique*. Celle de J. Denègre, qui évoque l'*objet géographique* et précise la forme de ces informations en fonctions de composantes gérées par les SIG, correspondrait quant à elle à l'étape de *formalisation technique*. La dernière enfin qui considère l'information géographique comme un ensemble de données concernerait plutôt l'étape d'*implémentation*.

En revanche, la définition informatique met en avant une autre sorte d'information. Elle présente des informations qui ne résultent pas de l'une des étapes du processus de conceptualisation, mais d'un autre processus qui irait en sens inverse (fig. A.1.6). Ce processus qui part des données pour aller vers les entités est le processus de traitement des données qui permet de créer de nouvelles informations ou connaissances sur les *objets géographiques*.

Dans cette thèse qui vise à produire plus d'informations sur les *objets géographiques* étudiés, nous travaillons essentiellement à la formalisation technique des *objets géographiques* en 'objets informatiques' ; nous développons une formalisation pour créer de nouvelles informations sur les 'objets informatiques' à partir de traitements sur les données pour permettre ensuite à l'utilisateur de disposer de plus d'informations à observer dans le SIG. Ce mouvement d'aller-retour entre *objet géographique*, 'objet informatique' et données, nous conduit suivant le contexte à employer le terme « information » dans un sens général (éléments de connaissance sur la réalité) ou dans un sens plus spécifique (informations issues du traitement de données). Par ailleurs, le terme *information géographique* qui

6. Cette adaptation pose le problème de savoir où introduire le caractère géographique. Par exemple : « faits et connaissances *(géographiques)* déduits des données *(géographiques)*. L'ordinateur manipule et gère des données. La signification déduite des données est l'information *(géographique)*»

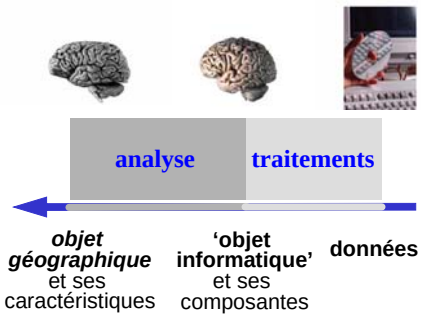


Figure A.1.6. *Le processus de traitements des données*

englobe plusieurs informations sur les *objets géographiques* (informations descriptives et information sur la localisation) est également fréquemment utilisé pour désigner l''objet informatique' dans un abus de langage courant dans le domaine des SIG. La figure A.1.7 récapitule l'ensemble des éléments terminologiques retenus.

Entité : « objet réel » (ex. : un pont)

objet géographique (en italique) :
résultat d'une formalisation par un thématicien (ex. : *pont*)

'objet informatique' (entre cote) :
résultat de la formalisation par un géomaticien (ex. : 'pont')

données : résultat de l'implémentation sous forme numérique.

Dans cette thèse suivant le contexte, le terme **information** évoquera :
- des **informations au sens général** décrivant des *objets géographiques* et 'objets informatiques' et
- des **informations au sens informatique**, c'est à dire le résultat de traitements sur les données.

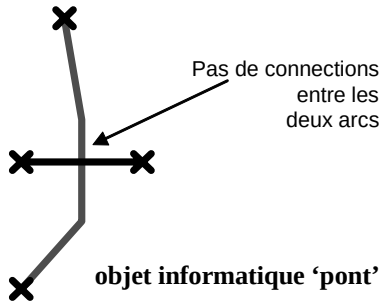
Par ailleurs, le qualificatif géographique associé à l'information, lui donne un sens particulier, dans le domaine des SIG :

information géographique sur un *objet géographique*
=
ensemble des composantes de l' 'objet informatique' (éléments décrivant les caractéristiques d'un *objet géographique* dont sa localisation, selon une modélisation propre aux SIG).

Dans cette thèse, nous tenterons de bien distinguer les différents termes. Néanmoins, il pourra nous arriver d'effectuer l'abus de langage fréquent dans le domaine des SIG, qui appose : **information (géographique) et 'objet'**, pour désigner l' 'objet informatique' avec ses composantes.
Par exemple, l'information 'route' (avec l'entre cote) signifiera la 'route' en tant qu' 'objet informatique' telle que décrite par ses composantes.

entité pont

objet géographique *pont*



objet informatique 'pont'

Figure A.1.7. *Eléments terminologiques*

A.1.2. SIG et informations géographiques

A.1.2.1. Les SIG comme SI pour les informations géographiques

L'utilisation de l'informatique pour gérer et étudier les *objets géographiques* s'appuie généralement sur un *système d'information géographique* ou SIG. Un système d'information géographique est en premier lieu un *système d'information (SI)*. Il a pour fonction de gérer, de traiter et de communiquer les informations qui lui sont confiées. La spécificité des informations géographiques a impliqué dès l'origine des outils particuliers. En effet, la caractéristique géographique nécessite des traitements appropriés à la fois pour sa gestion mais également pour sa communication. Or, les systèmes de gestion des SI traditionnels, fonctionnant à partir d'index alphabétiques ou ordinaux, ne disposaient pas d'index spatiaux leur permettant d'optimiser la recherche des informations à caractère spatial. De même, les fonctionnalités de restitutions des informations fournissaient des listes ou des tableaux, mais pas de cartes, mode privilégié de présentation des informations géographiques. Il a donc fallu développer des outils de gestion dédiés à ces informations.

A.1.2.2. Deux besoins à l'origine des SIG : le SIG au croisement de deux cultures

A l'origine des SIG comme pour toute application informatique, il y a le désir d'automatiser des tâches simples et régulières. En fait, deux besoins de nature très différente ont mené à l'élaboration des premiers SIG :

– un besoin de gestion de données comportant une composante cartographique, autrement dit le besoin de SGBDS ou L (système de gestion de bases de données spatialisées ou localisées) intégrant entre autre un index et des requêtes spatiales,

– un besoin de souplesse pour la représentation graphique d'informations, autrement dit le besoin d'outils de cartographie numérique adaptés à la représentation des multiples aspects sémantiques d'une même information géographique. Les outils de cartographiques numériques ne permettaient alors d'associer qu'une seule représentation à chaque objet.

Les SIG ont apporté une réponse unique, satisfaisant conjointement les deux besoins. Ils permettent non seulement de gérer l'ensemble des composantes de l'information géographique, mais également de les présenter et de les cartographier. Ainsi, depuis leur début les SIG procèdent de deux cultures : la culture informatique et la culture cartographique[7].

7. Cette position entre les deux cultures permet aux SIG d'occuper, encore aujourd'hui, une place particulière parmi les SI. Ils utilisent des outils et méthodes spécifiques tout en profitant des développements des deux domaines : informatique (ex. : géo-mining, introduction de programmes multi-agent, d'automates cellulaires) et cartographie (cartographie dynamique, cartographie du temps).

L'architecture de certains logiciels exprime cette dualité. Ces logiciels sont construits sur : d'une part un module de gestion des données de type SGBD classique et d'autre part un module de gestion des données graphiques. Ils induisent une vision « morcelée » de l'information géographique, où la dimension spatiale est dissociée de la dimension sémantique (cf. figure A.1.8). D'autres sont conçus autour d'un seul noyau-logiciel qui prend en charge les deux aspects, donnant à voir une information géographique plus unifiée.

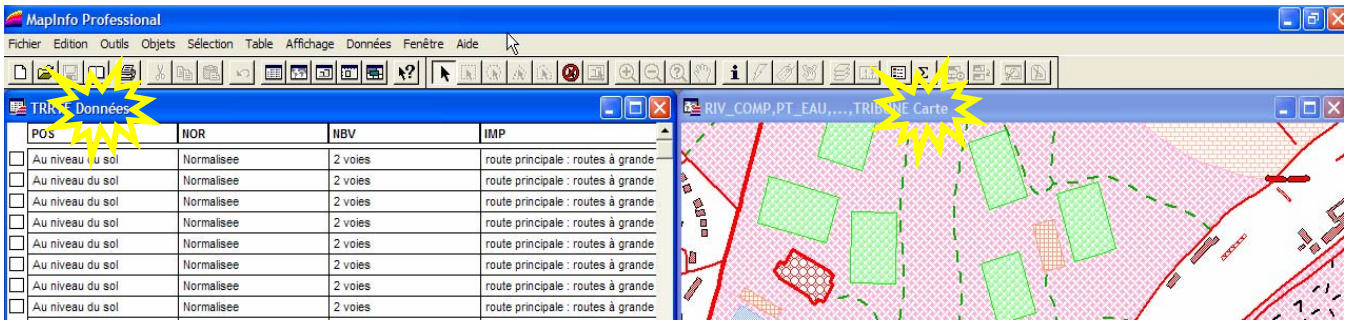


Figure A.1.8. *Une distinction entre les « données » et les « cartes » dans MapInfo*

Indépendamment de ces choix d'implémentation, de façon théorique, la plupart des logiciels-SIG conçoivent l'information géographique sur un *objet géographique* autour des trois même aspects :

- l'ensemble des caractéristiques sémantiques de l'objet (ex. : sa nature, ses propriétés),
- sa localisation,
- ses relations spatiales avec les autres objets.

En revanche, le modèle pour structurer les informations géographiques peut varier.

A.1.2.3. L'approche par champ et l'approche par objet

Il existe deux façons d'appréhender les informations géographiques : l'approche par champ (ou domaine ou zone [Bonin, 2002]) et l'approche par objet[8]. Elles conduisent à des modèles distincts de structuration en informations géographiques [Chrismann, 1975] et [Peuquet, 1986].

8. En anglais, on parle de *field approach* respectivement *object approach.*

A.1.2.3.1. L'approche par champ

Dans l'approche par champ, les objets sont considérés sous forme de distributions spatiales : à toute localisation est associée une valeur du *champ* considéré (ex. : champ de pesanteur). Leur représentation en informations géographiques associe alors une valeur d'attribut à toute localisation. Par exemple, une carte géologique ou une carte d'occupation des sols donne en tout point de l'espace une information sur la nature ou l'occupation des sols. Cette information peut être obtenue à l'aide d'une fonction mathématique interpolant la valeur en tout point à partir d'un ensemble d'observations (point, valeur observée). On obtient alors une représentation continue des informations (ex. : un champ d'altitudes) [Worboy, 1995] cité par [Li et al, 1999]. Mais, la représentation peut également être discontinue comme par exemple la représentation du relief par une hypsométrie qui associe une série de couleurs à une série de plages d'altitude. On notera que dans l'approche par champ, l'emprise au sol des faits géographiques n'existe pas explicitement. Les valeurs de l'information sont rattachées directement à la localisation. On peut donc trouver directement les valeurs en tout point. En revanche, il n'existe pas de relation entre les localisations possédant les mêmes valeurs. Aucun ensemble n'est défini. Ainsi, si l'on veut travailler sur les zones homogènes pour un attribut donné (ex. : les champs de blé, ou encore dans le cas de l'hypsométrie, les zones dont l'altitude est comprise entre 500 m et 550 m), ces zones doivent alors être reconstruites. En effet, elles n'existent pas explicitement en tant qu''objet', même si leur identification visuelle à partir d'une représentation graphique est immédiate.

A.1.2.3.2. L'approche par objet

Dans l'approche par objet, l'espace est considéré comme peuplé d'entités discrètes identifiables, localisées dans l'espace. Les entités sont décrites par des 'objets géographiques' caractérisés par un ensemble d'attributs et qui possèdent une emprise au sol. Cette emprise peut être décrite sous forme géométrique à l'aide d'objets simples (souvent des points, lignes ou polygones) et une localisation. L'entité peut également posséder des relations topologiques avec les autres entités. Ainsi, dans le cas de l'approche par objet, l'emprise au sol de chaque 'objet' est explicite ; Elle sert de support aux informations sémantiques associées. Elle fait le lien entre ces attributs et la localisation qui la caractérise.

Les deux approches utilisent des représentations différentes. Elles privilégient également certaines méthodes et outils d'analyses spatiales[9] (par exemple : les outils de géostatistiques qui permettent d'interpoler pour le modèle par champ, les méthodes d'analyse topologique ou d'analyse morphologique pour le modèle objet). Cependant, elles restent très complémentaires [Couclelis, 1992], [Engenhofer et Herring, 1991]. Cela explique les travaux visant à développer un modèle intégré dans les SIG que [Cova et al., 200] propose d'appeler Object-field (ou O-field) [Li et al, 1999].

9. « analyses spatiales » étant pris ici dans le sens de la géomatique d'« analyses mathématiques et / ou informatique comprenant un aspect spatial (analyses morphologique, topologique ou statistique). Dans le chapitre B.1, nous reviendrons sur ce terme et ferons le lien avec le terme d'« analyse spatial » en géographie.

A.1.2.4. Les composantes de l'information géographique

Conçue autour de trois aspects, l'information géographique sur un *objet géographique* se décompose généralement dans les SIG selon trois composantes : géométrique, sémantique et topologique.

A.1.2.4.1. La composante géométrique

La composante géométrique correspond à la représentation de l'emprise au sol de l'information géographique. Elle décrit :

- − la localisation de l'information, sous forme de coordonnées (x,y,(z)), par rapport à un système de référence (cf. figure A.1.9),
- − la morphologie de cette emprise, sa forme. Elle comprend des paramètres tels que la surface, le périmètre, etc.

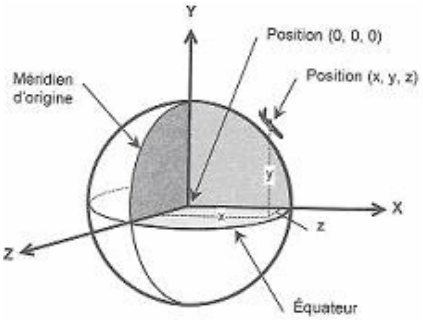


Figure A.1.9. *Exemple de coordonnées (x,y,z) dans un système de référence de type euclidien*

Il est important de souligner que la composante géométrique, si elle peut servir de support cartographique pour l'information géographique, ne se confond cependant pas avec la composante graphique. Par exemple, un bâtiment dont la composante géométrique est un polygone peut être représenté par un pictogramme exprimant sa fonction (ex. : ,). Le lien reste cependant très fort [Peuquet, 2002, p. V], [Langran, 1993], jusqu'à influencer les modes de mises à jour des bases de données géographiques.

Il existe deux modes de représentations informatiques pour implémenter la composante géométrique sous forme de données : le mode raster (ou matriciel) et le mode vecteur.

Les modes de représentation : matriciel ou raster

Le mode raster, aussi appelé matriciel, décompose l'image sous forme d'une matrice ou d'une grille et associe une valeur à chaque carré élémentaire − pixel (fig. A.1.10). Cette valeur peut être un niveau de gris, une couleur (comme dans les *orthophotographies* aériennes ou satellitaires, les cartes scannées). Elle peut aussi correspondre à une radiométrie en infrarouge ou radar ou encore à toute autre valeur comme une altitude ou un nombre d'habitants (on appellera ici ces images : des *images sémantiques)* (fig. A.1.11).

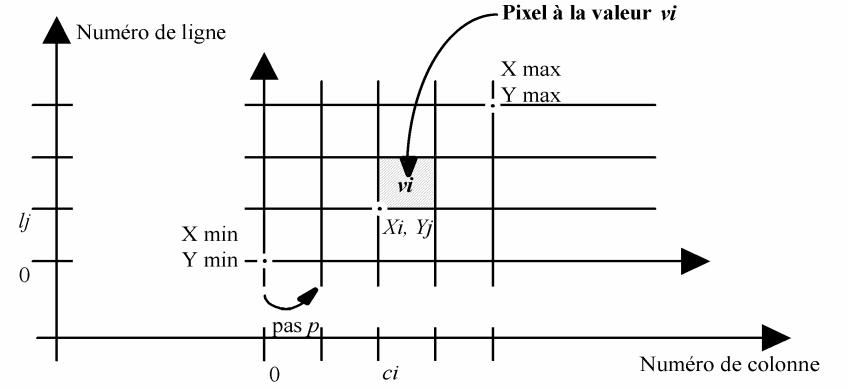


Figure A.1.10. *Principe d'une image raster*

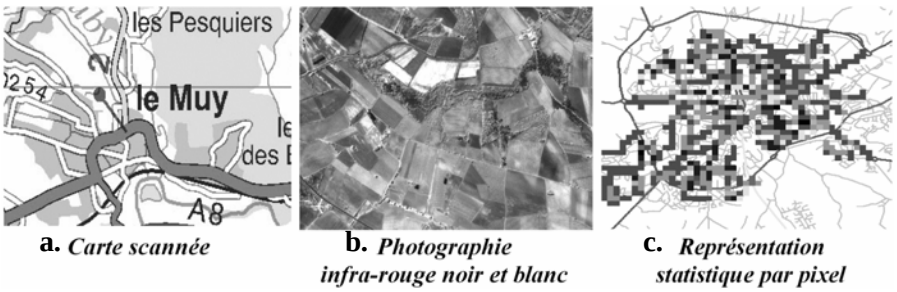


Figure A.1.11. *Exemples d'images raster*

Pour réaliser des études à partir d'informations raster, on utilise des fonctionnalités de traitement différentes selon que l'information est issue de sources radiométriques (fig. A.1.11b) ou de sources statistiques (fig. A.1.11c). Les fonctionnalités de traitements d'image servent à réaliser des études variées à partir d'orthophotographies (ex. : suivi des inondations, étude du développement urbain, bilan de catastrophes naturelles). Elles permettent d'identifier et d'extraire des 'objets géographiques' des photographies (ex. : identification de 'zones inondées', de nouveaux 'habitats', de 'forêts' décimées). Pour l'analyse des images raster sémantiques, la variable représentée sur chaque pixel correspondant à une variable thématique, on utilise alors principalement des traitements statistiques (ex. : étude de répartitions) [Latour et Le Floc'h, 2001]. Le cas des cartes scannées (fig. A.1.11a) est un cas un peu particulier. Bien qu'elles puissent représenter des informations sémantiques (ex. : cartes de la populations), elles sont de sources radiométriques (capture de la couleur par le scanner, la valeur de chaque pixel traduit une couleur et non à une valeur thématique). C'est pourquoi pour leur analyse, on utilise des fonctionnalités de traitements d'images, et le plus souvent il s'agit alors de reconstituer des informations en mode vecteur.

Le mode vecteur

Alors que le mode raster utilise une grille pour décomposer la représentation en cellules élémentaires, le mode vecteur la décompose en ses éléments constituants. Pour les expliciter, il s'appuie sur trois primitives géométriques : le point, la polyligne et le polygone (fig. A.1.12).

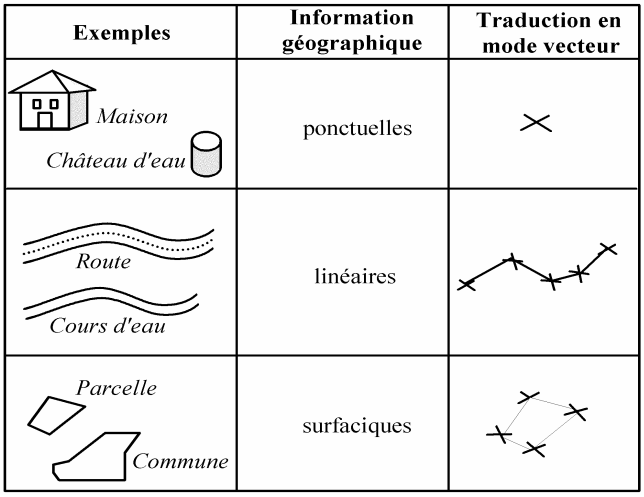


Figure A.1.12. *Les primitives géométriques du mode vecteur*

Un même *objet géographique* peut être représenté par des données différentes en mode vecteur (ex. : une route peut être représentée par une ligne ou par un ensemble de surface contiguës, une maison par un polygone ou par un point). Il s'agit ici de modélisations distinctes. La détermination de ce qui constitue la « bonne modélisation » revient au concepteur des données. Il choisira la façon de représenter chaque *objet* en fonction de sa «vision du monde»[10]. Il décidera également quels seront les *objets* représentés et ceux qui ne le seront pas. Les choix du concepteur ne se limitant pas à la forme, il décidera également de la composante sémantique de chaque 'objet', c'est à dire quelles informations attributaires viendront enrichir la description morphologique de l''objet'.

10. La « vision du monde » inclut également les contraintes auquel le concepteur des données peut être soumis : contrainte de coût, de délais, de faisabilité, d'accessibilité des données, etc.

Mode raster versus mode vecteur

Les deux modes de représentation des données géographiques sont complémentaires. Ils ont des qualités et des utilisations différentes. Certaines sont résumées dans la figure A.1.13.

Mode vecteur	Mode raster
Volume de stockage compact	Volume de stockage important
Structure de données complexe	Structure de données simple
Mise en œuvre facile de la topologie	Mise en œuvre difficile de la topologie
« Saisie » possible des objets	On ne «saisit » que des pixels
Ne contient que les informations spécifiées	Exhaustivité des photographies raster
Collecte des informations de mise à jour	Rapidité de la mise à jour
Croisement thématique plus complexe	Croisement thématique simple

Figure A.1.13. *Avantages respectifs de chaque mode de représentation*

Comme le souligne certains auteurs ([Bonin, 2002, p.13], [Couclelis, 1993]) il existe des rapprochements possibles entre les modèles de structuration de l'information géographique d'une part et leurs modes de représentation en données d'autre part :

- le mode vecteur qui associe à chaque information géographique une composante géométrique géoréférencée et une composante sémantique est proche du modèle objet,
- le mode raster qui attribue la valeur d'une variable descriptive en chaque point (défini par un pixel) de l'espace, est proche du modèle par champ. Ainsi, certains considèrent que le mode raster est un modèle par champ discret.

Cependant, ils ne se confondent pas. Ils ne relèvent pas du même niveau de conceptualisation. Ainsi les modèles de représentation s'appliquent à des données, tandis que les modèles de structuration concernent les informations. Pour nos recherches nous nous sommes placés dans le cadre d'une approche par objet, non seulement parce que nos travaux s'appuient essentiellement sur des données en mode vecteur, mais surtout parce que le concept d'objet (géographique et informatique) est un élément clé de notre méthode. En effet, nous utilisons des données, mais c'est avant tout pour construire une méthode de suivi de phénomènes.

A.1.2.4.2. La composante sémantique

La composante sémantique regroupe toutes les informations descriptives thématiques codées sous forme de caractères (ex. : nombre de bâtiments, sens de circulation, taux de pollution, nom, etc.), mais aussi d'images (ex. : photo de la façade), de sons, de films (cf. figure A.1.13). Par exemple, un bien immobilier peut être décrit par toute une liste d'éléments classiques (nombre de pièces, superficie, étage, etc.), mais également par un plan des pièces, une photo de sa façade, et même par une visite guidée sous forme de film. Il existe ainsi des bornes informatiques à usage touristique, qui donnent vocalement le nom des sites lorsque l'on balaye un lieu avec la souris[11].

11. Certains chercheurs travaillent sur des cartes pour aveugles, où les mots et autres sons seraient utilisés pour communiquer les informations exprimées habituellement sous forme graphique. Couplées à un GPS, elles pourraient aider à la navigation.

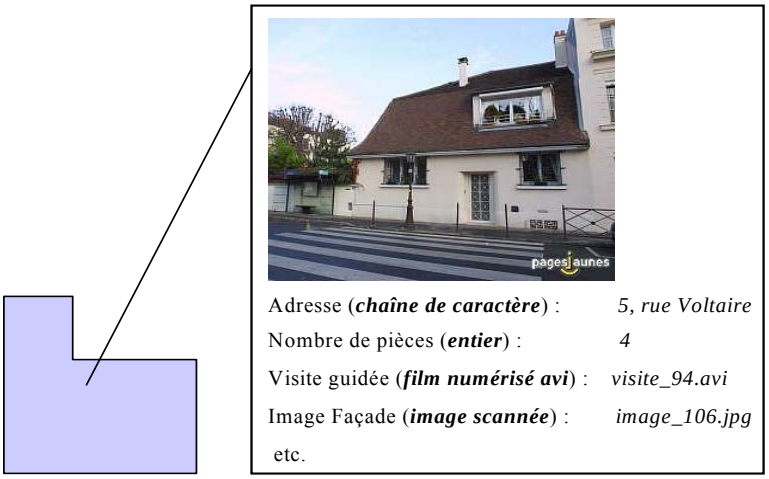

Adresse (chaîne de caractère) : 5, rue Voltaire
Nombre de pièces (entier) : 4
Visite guidée (film numérisé avi) : visite_94.avi
Image Façade (image scannée) : image_106.jpg
etc.

Figure A.1.14. *Exemples d'informations sémantiques*

La composante géométrique et la composante sémantique associées à chaque information géographique n'ont pas toutes la même importance relative. Il existe en effet des informations géographiques essentiellement sémantiques. Dans ce cas, la composante géométrique se réduit à un point tandis que la composante sémantique comporte un très grand nombre d'informations (par exemple, dans les DDE (Directions Départementales de l'Equipement) l'application-SIG gère les accidents sous forme de ponctuels associés à près de 300 attributs)[12]. A l'inverse certaines données sont essentiellement géométriques et contiennent peu d'informations sémantiques (c'est généralement le cas des données *généralistes supports,* comme la BD TOPO® de l'IGN, qui servent à recevoir ou à positionner d'autres informations plus thématiques). Très complémentaires dans l'analyse de phénomènes géographiques, elles ne font généralement pas appel aux mêmes outils d'analyses. Dans le paragraphe B.1.2.2, nous proposons une typologie pour les différencier en vue de formaliser leur intégration sous forme de données géographiques et permettre ainsi de mieux les exploiter.

12. On peut même envisager des informations géographiques sans composante géométrique. En effet, certains utilisateurs de données géographiques localisent leurs informations sous forme d'un nom de ville, d'un code INSEE, d'une adresse dans un tableau, mais ils n'effectuent pas l'opération de *géocodage*, qui permet d'associer des coordonnées aux informations pour les positionner explicitement sur une carte.

A.1.2.4.3. La composante topologique

La composante géométrique est une information de positionnement absolu lié à un référentiel, alors que la composante topologique est une information sur le positionnement relatif des objets entre eux. Elle regroupe les informations topologiques qui sont, par définition mathématique, des « informations invariantes par homéomorphisme qui entrent dans une analyse spatiale ». Plus concrètement, elle porte sur la géométrie, mais ne concerne que les propriétés comme l'intersection, l'inclusion, la connexion. Suivant les bases de données et les logiciels de SIG, cette composante, fondamentale pour toutes les applications de diffusion et de propagation, est codée explicitement ou recalculée à partir de la géométrie.

A.1.2.5. La modélisation des informations géographiques en base de données

La représentation des entités géographiques sous forme de données implique de nombreux choix, à la fois en ce qui concerne :

- la composante géométrique : quel mode de représentation sera retenu (ex. : raster ou vecteur, quelle échelle d'utilisation)? Les bâtiments en

mode vecteur, seront il représentés ? Si oui, sous quelle forme (point ou polygone) ?,

- la composante sémantique : quelles caractéristiques seront représentées (ex. : sur une route indiquera t'on le nombre de voies, son sens, sa largeur, sa vitesse maximale autorisée, nature du revêtement, etc.) ? Et sous quelles formes (ex. : liste de valeurs restreintes, nombre entier ou réel, etc.) ?

Ces choix qui doivent exprimer la « conception du monde » de l'observateur sous forme de données, participent à la fois de la formalisation conceptuelle et de la formalisation technique. Pour le producteur de base de données, leur élaboration constitue la *modélisation conceptuelle* et leurs définitions sont consignées dans le document appelé *spécifications de contenu*.

Suivant la personne qui réalise la *modélisation conceptuelle* (et suivant son point de vue), suivant ses compétences (plutôt thématiques ou plutôt techniques), le résultat de cette modélisation peut prendre des formes très différentes. Elle peut correspondre à des niveaux d'observation divers et conduire à des niveaux de représentations variées, plus ou moins détaillés (par exemple : la ville peut être vue comme une entité ponctuelle dans une représentation générale du territoire ou bien comme une agrégation d'autres entités géographiques plus petites comme des îlots voire des bâtiments, des foyers). Par ailleurs, au moment de rendre cette formalisation opérationnelle, elle peut être décrite de façon plus ou moins précise, avec des spécifications plus ou moins développées (par exemple : « les *villes ponctuelles* considérées ne pourront être que des villes de plus de 10 000 habitants, desservies par une infrastructure ferroviaire et située à moins de 10 km d'un échangeur d'autoroute, présentant une population active de plus de 20%. Elles devront être reliées systématiquement au réseau représentant les infrastructures de transport, etc. »). Le niveau de détails des spécifications du modèle évolue en fonction du niveau d'analyse du problème et du degré d'intégration des contraintes méthodologiques et techniques qui influent sur la solution. La modélisation est souvent un processus graduel récursif. Ainsi, certaines modélisations conceptuelles se limitent à sélectionner parmi les objets réels et leurs attributs ceux qui contribuent au phénomène étudié (il s'agit plus d'une formalisation thématique). D'autres anticipent déjà sur le processus de modélisation logique ; ils commencent à spécifier certains détails utiles en vue d'une implémentation ultérieure, pour s'assurer de pouvoir ensuite réaliser certaines analyses spécifiques. La modélisation conceptuelle présente alors les *objets géographiques* sous une forme proche de leur implémentation logicielle à venir (la modélisation intègre déjà la formalisation technique). Par exemple, si l'analyse implique de disposer d'une représentation en vraie 3D, pour structurer la composante géométrique des informations géographiques, on peut dès le départ anticiper le recours à un modèle différent de celui utilisé classiquement par les SIG. En effet, leur modèle conceptuel s'appuyant sur des objets 2D (points, lignes et polygones) ne peut pas générer de « vraie » 3D mais seulement des « fausses 3D » (dites parfois 2D ¼ ou 2D ½ pour spécifier leur niveau de complexité) [Poupeau et Bonin, 2006].

Pour ces travaux portant sur l'élaboration d'une méthode générique pour l'observation et le suivi de phénomènes géographiques dans un SIG, nous nous situons fondamentalement dans le cadre de la *modélisation conceptuelle* :

- là où les informations géographiques n'ont pas encore été traduites sous forme de données géographiques, ce qui ne fige pas les composantes géométriques et sémantiques,

- là où l'on peut encore raisonner sur les *objets géographiques* et chercher à articuler leurs diverses représentations thématiques en informations géographiques, ce qui nous permet d'envisager des représentations à différentes échelles (à des *niveaux de détails* de représentations différents).

Nous avons également adopté quelques préalables. En posant l'utilisation des SIG, nous avons opté implicitement pour le **mode vecteur**. Par ailleurs, nous souhaitions contribuer par ces recherches aux applications thématiques des SIG. Ainsi, sans que cela soit spécifié au départ, nous avons voulu faciliter la mise en œuvre ultérieure de ces travaux par des utilisateurs qui ne sont pas tous des spécialistes des SIG. Cela nous a conduit à travailler sur des **données existantes** et mises à jour de façon homogène sur l'ensemble du territoire (en particulier la BD TOPO® de l'IGN), sur des **logiciels SIG du marché** et nous le verrons par la suite sur une **modélisation du temps simple** et couramment utilisée (le *snapshot*). La partie B qui présente notre proposition de méthode de suivi de phénomènes géographiques, s'appuie sur ces choix préalables. Elle montre également comment intégrer des éléments de choix techniquement plus développés pour produire une méthode d'observation plus fine.

A.1.3. De l'analyse dans les SIG aux SIG outils d'aide à la décision

A.1.3.1. Les utilisations des SIG

Il existe aujourd'hui des systèmes de gestion de base de données (SGBD) capables de gérer une composante géométrique (ex. : Oracle avec spatial carthridge, PosgreSQL avec Posgis). Les applications-SIG impliquant des volumes de données importants leur sous-traitent d'ailleurs souvent la gestion de leurs données (ex. : les applications du géomarketing). Par ailleurs, certains traitements de rédaction cartographique professionnelle restent du ressort d'outils spécifiques (ex.: illustrator, freehand). Mais les SIG proposent d'autres fonctionnalités en plus des fonctionnalités de gestion et de représentation de la composante géométrique, en particulier des fonctionnalités d'analyses (par exemple : calcul de la moyenne d'un attribut sémantique, de la distance ou détermination de l'appartenance à une zone) [Pornon, 1990], [Denègre et Salgé, 1996], [Bordin, 2002]. Très variées, elles concernent l'ensemble des composantes de l'information géographique. Les fonctionnalités d'analyses standards participent essentiellement à des analyses descriptives, à la création de nouveaux objets ou aux vérifications d'intégrité de la base. Cependant la possibilité de combiner des fonctionnalités élémentaires ou d'en développer de nouvelles permet aux utilisateurs de réaliser des analyses plus complexes. En fait, les fonctionnalités d'analyses occupent une place plus ou moins importantes selon les applications et le type d'usage des SIG.

[CODATSI, 1996] et [Georget, 2000] distinguent trois types d'usage des SIG[13]:

- le *SIG-inventaire/observatoire* a pour objet de répondre à des attentes en termes de connaissances sur un territoire. Il utilise principalement des fonctionnalités de gestion pour répondre aux requêtes simples de types Où ? et Quoi ? portant respectivement sur la composante sémantique (ex. : où sont les écoles maternelles? où est la rue Guillebaud ?) et sur la composante géométrique (ex. : « quoi » est là? quel type d'objet possède cette emprise? Quelles sont ses caractéristiques ?) [Denègre et Salgé, 1996].

13. Dans [Bordin, 2002], nous suggérons d'en ajouter un quatrième, complémentaire aux trois proposés: le *SIG-communication* qui sert à la diffusion et la transmission d'informations géographiques. Nous ne le prendrons pas en compte dans ces travaux car il est plus du registre de la présentation des informations géographiques que de leur étude.

– le *SIG-étude/analyse* a pour objet de mettre en évidence des faits spatialisés, de réaliser des analyses, de comparer des scénarios, etc. Il peut utiliser des fonctionnalités d'analyses diverses et les combiner en fonction de l'application (ex. : analyse topologique pour une optimisation d'itinéraire, analyse statistique sur des valeurs d'attributs sémantiques, calculs des intersections de deux partitions du territoire)

– le *SIG-gestion* exploite des informations géographiques dans le cadre de procédures préétablies, en vue d'une meilleure gestion des objets géographiques décrits (ex. : gestion du PLU, des accidents, de la sectorisation scolaire, des candélabres). Comme pour le SIG-étude, le SIG-gestion fait appel à des analyses variées. La différence réside en ce que ces analyses sont prédéfinies et formatées dans des processus stables. Contrairement au SIG-analyse qui crée des informations nouvelles, le SIG-gestion met à jour une information élaborée auparavant. L'application SIG cherche à atteindre un rythme de croisière.

Ces divers usages, non exclusifs les uns des autres, se succèdent fréquemment dans le temps d'un projet de SIG ou dans l'étude d'un phénomène localisé. Lorsque le SIG-gestion est tout à fait opérationnel, c'est souvent l'occasion pour l'utilisateur d'observer les anomalies ou de s'intéresser à des problèmes nouveaux, ce qui le conduit alors à développer un autre projet de SIG d'observation et d'analyse (fig. A.1.15).

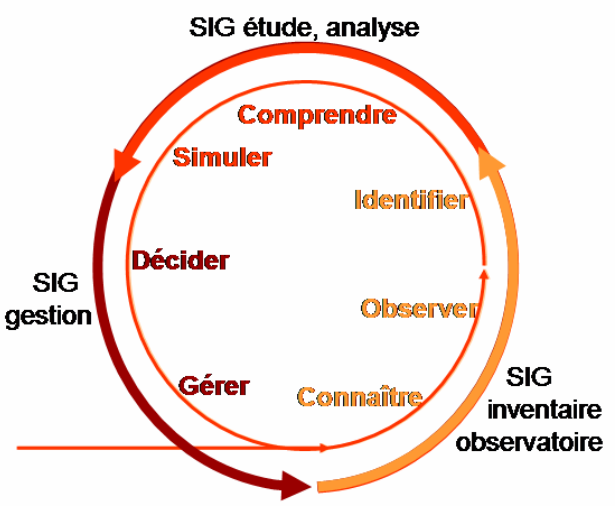


Figure A.1.15. *Cycle des différents usages des SIG.*

Pour celui qui met en place une solution SIG, ces différents types correspondent à la fois à une plus grande maîtrise des fonctionnalités des SIG (des requêtes simples vers le développement d'applications) et à une meilleure connaissance du territoire. Les trois usages concordent avec les étapes d'un processus d'étude qui irait de la connaissance du territoire à une identification plus fine de ce qui s'y joue en vue d'identifier des phénomènes, de les comprendre pour tenter de les simuler et ainsi prendre des décisions, pour mieux les gérer (cf. l'intérieur de la figure A.1.14). Le suivi des effets et leur observation dans un nouveau cycle d'analyse permettant par feed-back de prendre en compte de nouvelles données et d'améliorer le processus de décision.

A.1.3.2. Les SIG-inventaire/observatoire comme préalables aux analyses dans les SIG

Il est intéressant de noter ici combien l'usage des SIG-inventaire/observatoire a changé depuis l'origine. Les premiers SIG servaient aux organismes à recenser et à gérer leurs données géographiques (« Quelles *données* possédons-nous et où sont-elles ? ») ; ils jouaient le rôle de « meubles à plan ». Peu à peu, le besoin d'inventorier les données éparpillées dans les services a évolué vers le besoin d'inventorier les objets géographiques du territoire[14] (« Quels *objets* se trouvent sur notre territoire et où sont-ils ? »). La localisation des informations géographiques dans les SIG a permis l'étude de leur répartition spatiale, l'identification de polarisations ou de concentrations, la mise en évidence de phénomènes spatiaux (telle que la concentration de la majorité des cas de saturnisme sur une moitié de Paris). Avec ces premières applications, les utilisateurs ont acquis une meilleure représentation de leur(s) territoire(s) et de les *objets* géographiques présents. Ils ont parallèlement accumulé des données à l'occasion des mises à jour rendant envisageable aujourd'hui la prise en compte de la dimension temporelle.

14. Une rencontre en 2001 avec un utilisateur illustre ce point. Cet utilisateur, en phase initiale de projet SIG, voulait améliorer la gestion de son cadastre. Accessoirement, il espérait inventorier les poubelles publiques de sa commune pour mieux gérer la collecte des déchets, car il n'avait aucune idée du nombre exact de poubelles, ni de leur localisation. En effet, la décision d'en installer une, était souvent prise à l'occasion de la réfection d'un trottoir ou d'un carrefour, sans s'inscrire dans un projet de gestion des espaces publiques ou de toute autre planification.

De nouveaux besoins apparaissent. Les utilisateurs à la recherche d'une plus grande connaissance de leurs territoires tentent d'extraire de leurs données des informations nouvelles susceptibles de les aider à prendre leurs décisions. L'observation de la dimension temporelle leur semble prometteuse. Cependant, les difficultés pour l'introduire dans les bases de données, dans les logiciels et les raisonnements, n'ont pas encore permis d'aboutir à des solutions simples clés en mains. Beaucoup d'analyses spatio-temporelles s'effectuent aujourd'hui par simple comparaison visuelle de deux états de la base de données, le SIG redevenant à cette occasion simple outil d'inventaire cartographique (fig. A.1.16).

Figure A.1.16. *L'analyse spatio-temporelle visuelle à partir de deux versions différentes de la BD TOPO®*

Il existe donc un réel besoin de méthodes et d'outils pour identifier et suivre des phénomènes spatio-temporels. Dès lors l'enjeu des SIG-observatoire devient l'observation des territoires dans l'espace *et* dans le temps avec entre autre pour objectif de fournir des informations aux SIG-analyse et d'aider à la décision [Ieti.fr], [De Blomac, 2004]. Or pour fournir des informations aujourd'hui, les SIG-Observatoire requièrent souvent la participation visuelle et intellectuelle de leurs utilisateurs. Dans ce contexte, proposer une méthode générique pour construire de tels indicateurs participerait au développement des SIG et permettrait aux utilisateurs d'aller plus loin dans l'observation des phénomènes spatio-temporels. Cela contribuerait également au développement des SIG en tant qu'outils d'aide à la décision. En effet, si les SIG interviennent déjà dans les processus de

décision, ils ne sont pas réellement des systèmes d'aide à la décision (SIAD).

A.1.3.3. Vers un SIG, outils d'aide à la décision

Les SIG en tant qu'outils d'analyse semblent pouvoir s'inscrire dans une famille d'outils informatiques plus large : les *systèmes informatisés d'aide à la décision* (ou SIAD). Les SIG pourraient alors bénéficier de leurs avancées méthodologiques et techniques. Ainsi, lorsque les outils de gestion de bases de données relationnelles classiques ont été capables d'intégré la dimension spatiale (ex. : Oracle avec sa cartouche spatiale), les SIG ont tiré profit de leur capacité de traitements, en particulier en terme de volume de données, d'intégrité et de sécurité. Cependant, si le problème originel de gestion des données géographiques est pris en charge aujourd'hui par des outils informatiques généralistes, d'autres besoins d'analyse ou de représentations de ces informations, restent encore spécialisés. A l'inverse, les outils d'aide à la décision présentent des capacités qu'il serait utile de développer dans les SIG.

Les SIAD appartiennent au domaine du *Business Intelligence* (BI)[15], c'est à dire de l'informatique appliquée à l'aide à la décision. Ils participent aux outils décisionnels dont le but est *« l'exploitation des données des entreprises afin de faciliter la prise de décision par les décideurs par une meilleure compréhension du fonctionnement de l'organisation et une anticipation des actions par un pilotage éclairé de cette organisation »* [ccm.fr]. La structure générale de ces outils (fig. A.1.17) fait apparaître trois fonctions principales :

- abstraction, intégration (fig.)
- gestion, archivage (fig.)
- analyse, distribution (fig.)

15. ou informatique décisionnelle ou encore intelligence d'affaire. Notion apparue dans les années 70 avec les premiers info-centres (systèmes qui envoyaient des requêtes directement aux serveurs de production). Dans les années 80, le mode client serveur a permis d'isoler la production de données et les dispositifs décisionnels. Dans la foulée, des acteurs spécialisés se sont lancés dans la définition de couche d'analyse « métiers » avec des masques de données. Aujourd'hui les BI ne sont plus seulement réservés aux équipes techniques, ils sont aussi accessibles aux responsables opérationnels.

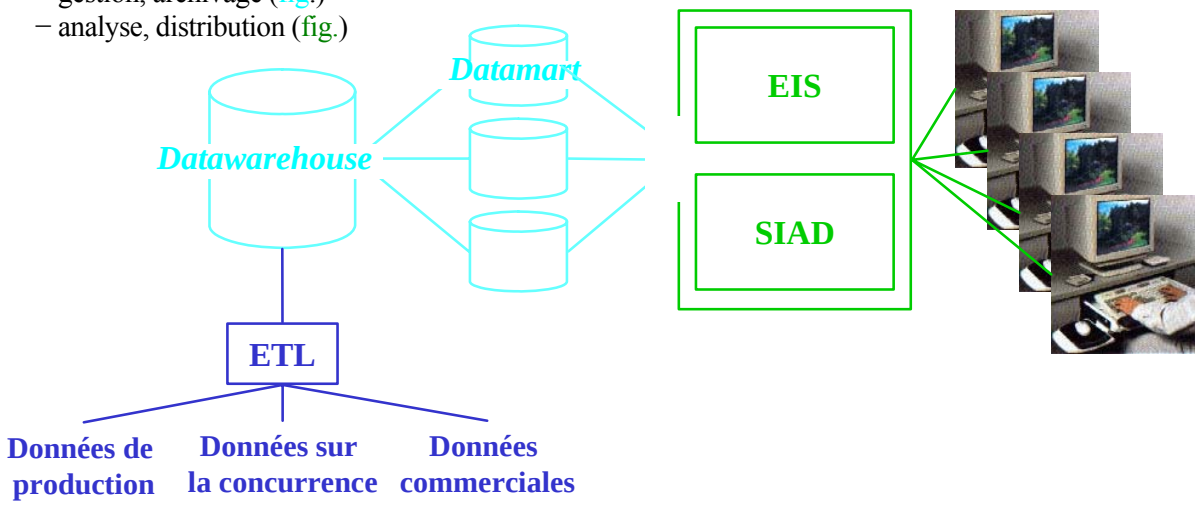


Figure A.1.17. *Le SIAD au sein de l'outil décisionnel*

ETL (extract, transform and load) : outil chargé d'extraire les données des différentes sources de les nettoyer, de les charger dans l'entrepôt de données ou **datawarehouse**, sorte de collection de données orientées sujet (i.e. : triées par thèmes), intégrées (nettoyées et normalisées), non volatiles (une fois dans le système elles n'ont pas vocation à sortir), historisées (datées). Ainsi défini le concept de datawarehouse semble proche de l'ensemble des données que peut contenir un SIG sur un territoire, classé souvent par thème ou couche (ex. : réseau routier, occupation du sol, administratif, population, etc.)

Datamarts ou magasins de données : sous-ensembles du datawarehouse pour un secteur particulier, sorte d'équivalents de ce que l'on appelle en géomatique des vues.

EIS (executive information system) :outil permettant d'organiser, d'analyser et de mettre en forme des indicateurs afin de constituer des tableaux de bords. Ce type d'outil, facile à utiliser, ne permet de manipuler que des requêtes préalablement modélisées par le concepteur.

SIAD (système informatisé d'aide à la décision) :système dont le but est de permettre la modélisation de représentations multidimensionnelle variées.

Il existe deux types d'analyse dans un outil décisionnel:

- *l'analyse multidimensionnelle* (AMD), qui consiste à modéliser des données selon plusieurs axes (ex. : chiffre d'affaire par type de clients, gamme de produit et zone géographique). Elle permet aux décideurs d'avoir accès rapidement et de manière interactive à une information pertinente, sous des angles multiples et divers, selon leurs besoins. Elle s'effectue dans l'EIS (executive information system)

- *l'analyse prédictive* qui exploite un ensemble d'évènements observés et historicisés, pour tenter de prévoir l'évolution d'une activité en dessinant des courbes de projection. Par exemple cette méthode peut s'appliquer à la gestion des relations clients, pour prédire leurs comportements, avec pour objectif de déterminer des profils d'individus présentant une probabilité importante d'achat ou pour prédire à partir de quand le client risque d'être infidèle. L'analyse prédictive utilise les outils de *Datamining*, technique d'analyse utilisant un logiciel pour « dénicher » des tendances ou des corrélations ou encore pour détecter des informations stratégiques ou découvrir de nouvelles connaissances en s'appuyant sur des méthodes de traitement statistique. Elle s'effectue dans le SIAD (système informatisé d'aide à la décision).

Des recherches en SIG existent pour développer des outils de *géomining*. Certains chercheurs travaillent par exemple à l'aide d'automates cellulaires [Engelen, 2005], [Ourednik, 2005], [Corcuff, 2005]). Il s'agit là principalement de travaux en mode raster, la structuration, en unités géométriques élémentaires régulières, de ce mode de représentation étant bien adaptée aux analyses statistiques. En revanche, les recherches à partir du mode vecteur, mode qui permet de continuer à travailler sur les objets géographiques, sont plus rares [Shiyuan et Deren, 2004].

Les SIG semblent en fait plus proches des outils décisionnels que de leur composante SIAD, car ils proposent comme eux un ensemble intégré de fonctionnalités (de l'acquisition à la restitution en passant par l'intégration, l'archivage et l'analyse). Construits pareillement autour d'un système de gestion, alimentés en données concernant un territoire (au lieu d'une organisation), les SIG permettent aux décideurs de sélectionner les informations géographiques selon plusieurs critères (géométriques, sémantiques et/ou topologiques). Ils donnent accès de manière interactive, à une information pertinente, présente sous des angles multiples selon les besoins. En revanche, ils peinent encore à détecter des informations stratégiques à partir d'informations archivées. Plus encore, ils ont du mal à exploiter la dimension temporelle des informations qu'ils gèrent, ce qui restreint de fait leurs capacités à déduire des informations nouvelles, à effectuer des analyses et des prévisions, et limite leur usage comme *outils d'aide à la décision*.

A.1.4. Intégration de la dimension temporelle dans l'observation des informations géographiques

Il existe des SIG -d'analyse sophistiqués. Toutefois, lorsqu'il s'agit d'informations géographiques - *temporelles*, l'usage des SIG se rapproche souvent du SIG -d'inventaire spatio-temporel visuel. Ainsi, la spécificité des informations géographiques différencie là encore les SIG(AD) des autres SIAD. En particulier concernant leur capacité à découvrir des tendances, des corrélations ou encore de nouvelles connaissances, ils manquent encore des outils pour le faire (géomining) et plus

fondamentalement, ils ne parviennent pas à tirer pleinement parti de la dimension temporelle. Or, le temps est une dimension importante :

- pour observer les phénomènes géographiques : il est à la fois un paramètre et une dimension d'observation. Ainsi, on peut chercher à déterminer la durée des phénomènes, mais aussi la date de leur survenue.

- pour analyser les phénomènes géographiques et prendre des décisions les concernant : Il aide à saisir et comprendre les évolutions, pour tenter de les prévoir, pour suivre les effets d'une décision, pour les modifier, etc.

Dans cette thèse, nous proposons une solution méthodologique pour mieux intégrer la dimension temporelle dans l'observation des phénomènes avec un SIG : Nous élaborons une méthode pour construire des indicateurs de suivi.

Dans une perspective plus large, nous imaginons alors que
− ces *indicateurs*, éléments d'analyse descriptive,
puissent également devenir
− des *paramètres* dans des processus d'analyse plus dynamiques, pour servir ensuite
− de *variables* dans des analyses de type géomining.

Ainsi, la méthode d'observation présentée dans ces travaux pourrait offrir des éléments pour répondre au besoin émergent de SIG –d'analyse spatio-temporelle et contribuer au développement d'outils d'aide à la décision géographique. Pour la construire, nous sommes partis du besoin en observation de phénomènes géographiques (chapitre A.2) et du bilan des moyens disponibles dans les SIG (chapitre A.3). Ces études préalables ont ensuite permis de spécifier l'apport de la méthode d'observation proposée.

A.2. Etudes de phénomènes géographiques

A.2.1. Observation de phénomènes

A.2.1.1. Phénomènes: éléments de définition

Pour observer un phénomène, on s'intéresse souvent à ses effets. On tente de comprendre le ou les processus sous-jacents en étudiant le mouvement et les évolutions des objets sur lesquels il opère. *Mouvement, évolution*, *processus* attestent d'une relation des phénomènes à la fois au temps et à l'espace. Toutefois, cette relation peut être envisagée de façon minimale, où le temps et l'espace n'interviennent que pour situer le phénomène comme un fait plus ou moins statique [Brunet et al, 1992]), comme *« ce qui se manifeste à la conscience par l'intermédiaire des sens…»* [Petit Robert, 2001], et dans un sens plus général, comme *« faits constatés qui constituent la matière de la science »* [Lalande, 1997]. Cependant, [Durand-Dastès, 2005] relève qu'il existe une autre conception qui considère le phénomène comme doté de dynamisme, où temps et espace procèdent du phénomène lui-même *«things occur and exist in relation to space and time »* [Peuquet, 2002, p. 11]. Il se définit alors comme *« un objet (d'expérience possible) qui apparaît dans l'espace et dans le temps »* [Petit Robert, 2001]. En posant l'observation de phénomènes géographiques, nous nous situons dans cette perspective où l'espace et le temps participent également aux phénomènes.

A.2.1.2. L'approche continue, l'approche discrète

L'observation de phénomènes, comme toute analyse, implique une étape de modélisation préalable. L'approche adoptée et le modèle utilisé diffèrent suivant le rôle donné au temps et à l'espace.

Partant de la distinction entre fait et phénomène effectuée par Paul Janet cité par [Lalande, 1997] *: « un fait est un phénomène arrêté, précis et déterminé, ayant des contours que l'on peut saisir et dessiner : il implique une sorte de fixité et de stabilité relative. Le phénomène c'est le fait en mouvement, c'est le passage d'un fait à une autre qui se transforme d'instant en instant »*. Nous distinguons deux façons de décrire les phénomènes :

- une *approche continue* principalement tournée vers le mouvement ;
- et une *approche discrète* fondée sur les faits.

L'**approche continue** considère le phénomène comme un mouvement, où *«…il n'y a pas de faits précis et déterminés, il n'y a que des nuances, c'est à dire des passages insensibles d'un phénomène à un autre »* (P.Janet, cité par {Durand-Dastès, 2005}). Les descriptions continues saisissent la forme du phénomène dans l'espace ET dans le temps. Il s'agit souvent de descriptions sous forme de formules mathématiques, comme par exemple celles utilisées pour les phénomènes en physique ou encore en chimie. La figure A.2.2 montre une description continue du phénomène de diffusion de l'innovation. La modélisation mathématique qui exprime le nombre de récepteurs touchés par l'innovation en fonction du temps a *« la forme bien connue des fonctions de type logistique »* [Saint-Julien, 2001].

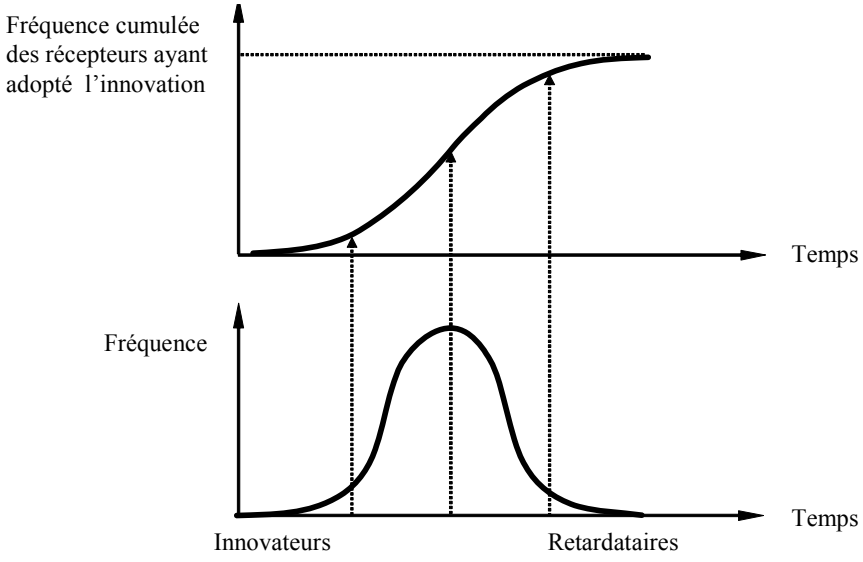


Figure A.2.2. *Les grandes étapes de la diffusion spatiale représentée comme un mouvement continu [Extrait de [Saint-Julien, 2001, p.167]*

Cependant selon [Pastor, 2004], l'utilisation de l'informatique implique une représentation discrète. Et plus généralement, comme le caractère continu veut qu'entre deux observations on puisse toujours en intercaler une troisième, il semblerait que l'observation d'un phénomène ne puisse être intrinsèquement que discrète (il a été montré que le regard n'est pas continu, les enregistrements de mesures non plus, et les capacités de traitement et de stockage d'informations ne sont pas infinies).

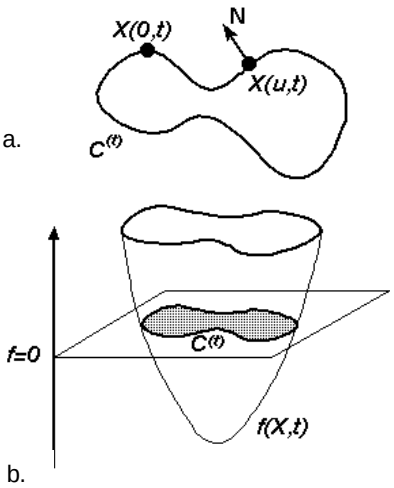


Figure A.2.1. *Représentation continue de la déformation d'une courbe 2D dans le temps*

Pour décrire une courbe C(t) qui se déforme au cours du temps [www-eph.int-evry.fr] propose deux types de représentations :

- Une représentation explicite où à tout instant t, la courbe C(t) est décrite par les coordonnées de ses points en fonction d'un paramètre u indépendant du temps (fig. a)

$$C(t)=\{X / X=X(u,t)\}$$

Les caractéristiques géométriques de la courbe s'expriment en fonction des dérivées partielles de X par rapport à u.
Dans un contexte discret, une représentation explicite implique un échantillonnage de la courbe.

- Une représentation implicite où à tout instant t, la courbe C(t) est décrite comme la ligne de niveau zéro d'une fonction scalaire f suffisamment régulière (fig. b):

$$C(t)=\{X / f(X(t)) = 0\}$$

L'**approche discrète** présente les phénomènes comme une succession de *faits*, distincts les uns des autres, décrits individuellement (fig. A.2.3). La série des mises à jour d'une carte donne un exemple d'une telle description. La figure A.2.4 présente ainsi l'évolution de l'occupation des sols entre 1982 et 1990, à partir de trois représentations d'états du territoire obtenues à partir de la base MOS (Mode d'Occupation des Sol)[2] de l'IAURIF.

Figure A.2.4. *Le phénomène représenté comme une successions de faits : l'exemple de l'occupation des sols représentée par une série de cartes*

La définition de Janet peut s'entendre avec des *faits temporels* (qui concernent des *événements*[3]). Cependant, la présentation générale s'appuiera sur des *faits spatiaux* (appelés *états*) comme dans l'exemple précédent (où le fait observé est l'occupation des sols). Elle reflétera ainsi mieux le cadre spécifique de ces travaux, car contraints par l'utilisation des SIG, nous ne travaillerons effectivement que sur de tels faits.

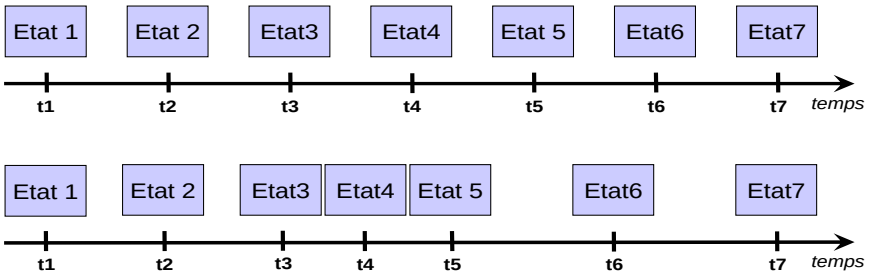


Figure A.2.5. *Rythme régulier ou irrégulier pour la description des faits (ou états)*

Dans l'approche discrète, les *faits* peuvent être décrits à des intervalles réguliers (correspondant par exemples à des contraintes techniques). Le rythme peut aussi être irrégulier pour s'adapter à la temporalité du phénomène (fig. A.2.5).

Dans cette description la restitution du mouvement est implicite : il se présente sous forme d'une succession de faits. **L'observateur doit lui-même effectuer le travail d'enchaînement** à partir de la série des états successifs. Il est souvent obligé de le faire visuellement. C'est le principe du cinématographe (cf. figure A.2.6). Il existe toutefois une description discrète plus explicite qui fournit directement à l'utilisateur les enchaînements ou *mutations*.

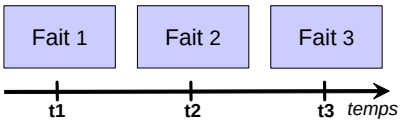


Figure A.2.3. *Représentation discrète sous forme de succession d'états*

2. Le MOS est la base de données d'occupation des sols réalisée et régulièrement mise à jour par l'IAURIF (Institut d'Aménagement et d'Urbanisme de la Région Ile de France). Nous l'avons exploitée pour mettre en œuvre nos travaux. Elle illustrera ainsi un certain nombre de points tout au long de cette thèse.

3. Nous définissions un *événement* comme un objet temporel. Il se caractérise par :

- sa localisation (dans le temps),
- son emprise temporelle (sa durée),
- ses caractéristiques sémantiques (ex. : sa nature, son amplitude, son nom, etc.),
- sa localisation spatiale

Il se présente comme le pendant dans *l'espace temps* des *objets* dans *l'espace géographique*. Par exemple, l'événement caractérisé par :

- date : *2005*
- emprise temporelle : *[24/8/05 ; 1/9/05]*
- caractéristique sémantiques : type *cyclone* ; nom : *Katrina* ; catégorie : *5*
- localisation spatiale : *Louisiane, Mississippi, Alabama et ouest des Etats-Unis*.

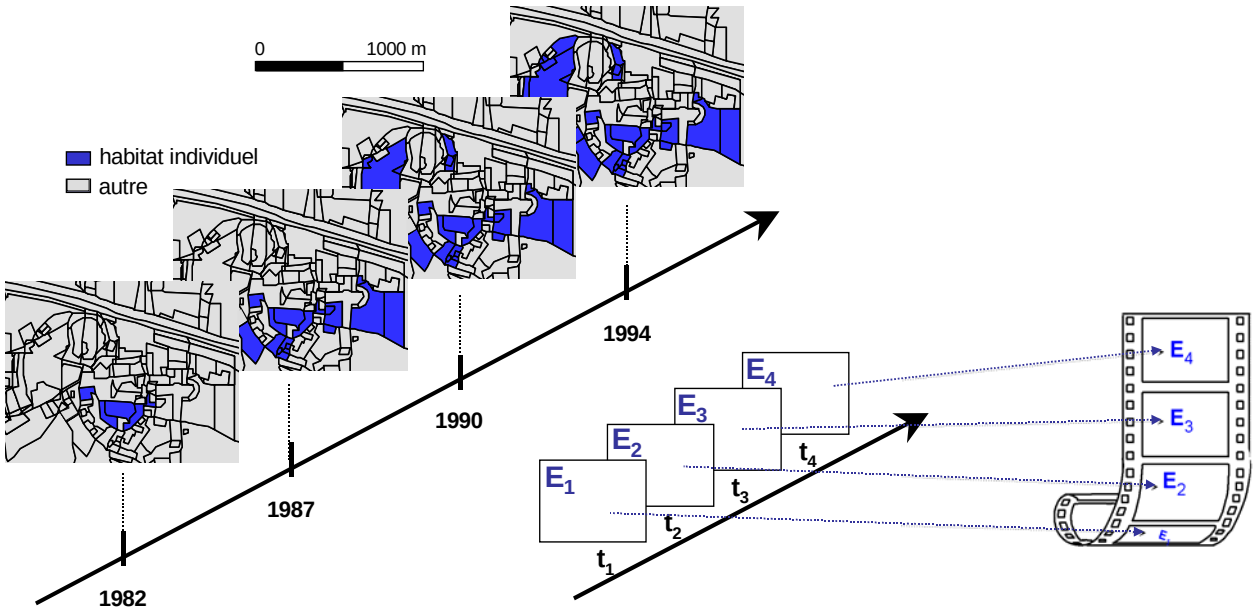


Figure A.2.6. *Le film, une représentation discrète du mouvement sous forme d'une succession d'états*

A.2.1.3. Représentation discrète du mouvement : suivi d'évolutions à l'aide de mutations

La représentation discrète par *mutations*[4] consiste à introduire explicitement les informations concernant les changements entre deux faits (cf. figure A.2.7).

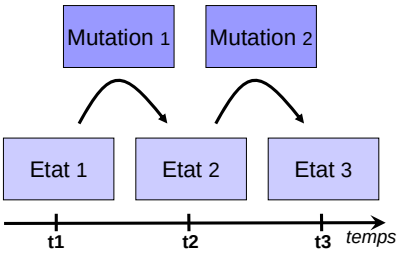


Figure A.2.7. *Le phénomène représenté par la série des mutations opérants sur les états*

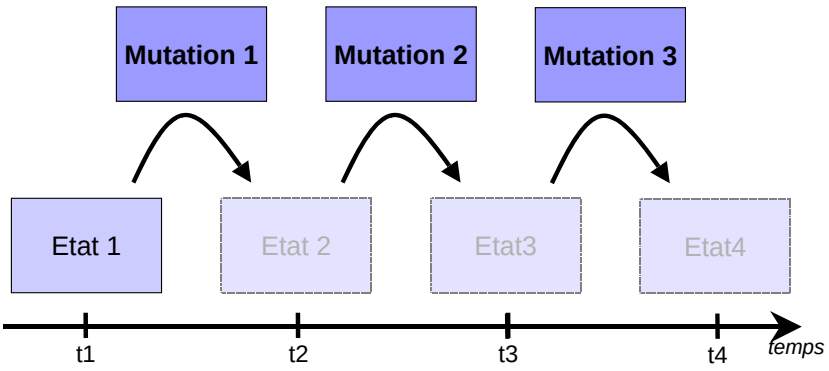


Figure A.2.8. *Le phénomène décrit par la série des mutations opérant sur les états*

La donnée de l'état initial et la suite des mutations suffisent pour reconstruire le mouvement : connaissant l'état 1 en appliquant la mutation 1 on obtient l'état 2, connaissant l'état 2 on en déduit alors l'état 3, et ainsi de suite (fig. A.2.8). En revanche, on ne peut pas accéder directement à un état. Il faut au préalable appliquer la suite des mutations. Cela explique pourquoi, on préfère souvent utiliser conjointement les deux descriptions discrètes et travailler simultanément avec l'ensemble des mutations *et* des états. La figure A.2.9 montre la complémentarité de ces deux descriptions. Cette figure qui présente l'évolution de l'habitat individuel comprend deux parties :

4. Une *mutation* n'est pas un *événement*. L'événement est la *cause* ou l'origine de changements produits sur les états. La mutation peut être *l'action* qui établit ces changements ou le *résultat* de cette action. Ainsi, l'événement peut engendrer des mutations, mais il ne se confond pas avec elles

– La partie inférieure correspond à une *description discrète sous forme de succession d'états*. Sur la portion de territoire étudiée, elle permet de :

- faire l'inventaire des zones d'habitat individuel
- constater que ce type d'urbanisation est un phénomène en développement sur les quatre périodes étudiées (la surface d'habitat individuel augmente sur chacune d'entre elles)
- constater que le phénomène tend à se diffuser autour des zones préexistantes.

Pour réaliser cette représentation, nous avons sélectionné parmi l'ensemble des thèmes contenus dans le MOS, le thème *habitat individuel* et nous avons affiché les quatre extraits correspondant aux quatre dates différentes.

– La partie supérieure correspond à une *description discrète par mutation*. Sur la portion de territoire étudiée, elle permet de :

- localiser les changements et de connaître leur nature,
- constater que le phénomène se développe de façon régulière (à chaque période de nouvelles zones apparaissent),
- constater que le type des mutations varie d'une période à l'autre.

Pour obtenir cette description, nous avons sélectionné les parcelles qui changeaient d'occupations des sols entre deux dates. Parmi elles, nous avons sélectionné celles qui se transformaient en *zone d'habitat individuel* et nous les avons affichées avec une légende correspondant au type d'occupation des sols avant leur mutation. Il importe de signaler ici la contribution de la base (le MOS) à cette analyse ; elle offre une représentation directe du phénomène à étudier (l'urbanisation d'habitat individuel) et elle facilite le suivi des zones dans le temps, à travers les états. L'utilisateur dispose rarement de conditions aussi favorables (ou bien les représentations des faits décrivant le phénomène ne sont pas disponibles, ou bien il n'y a pas de lien entre ces représentations permettant de les suivre au travers des états). L'un des enjeux de cette thèse sera de lui procurer des conditions équivalentes.

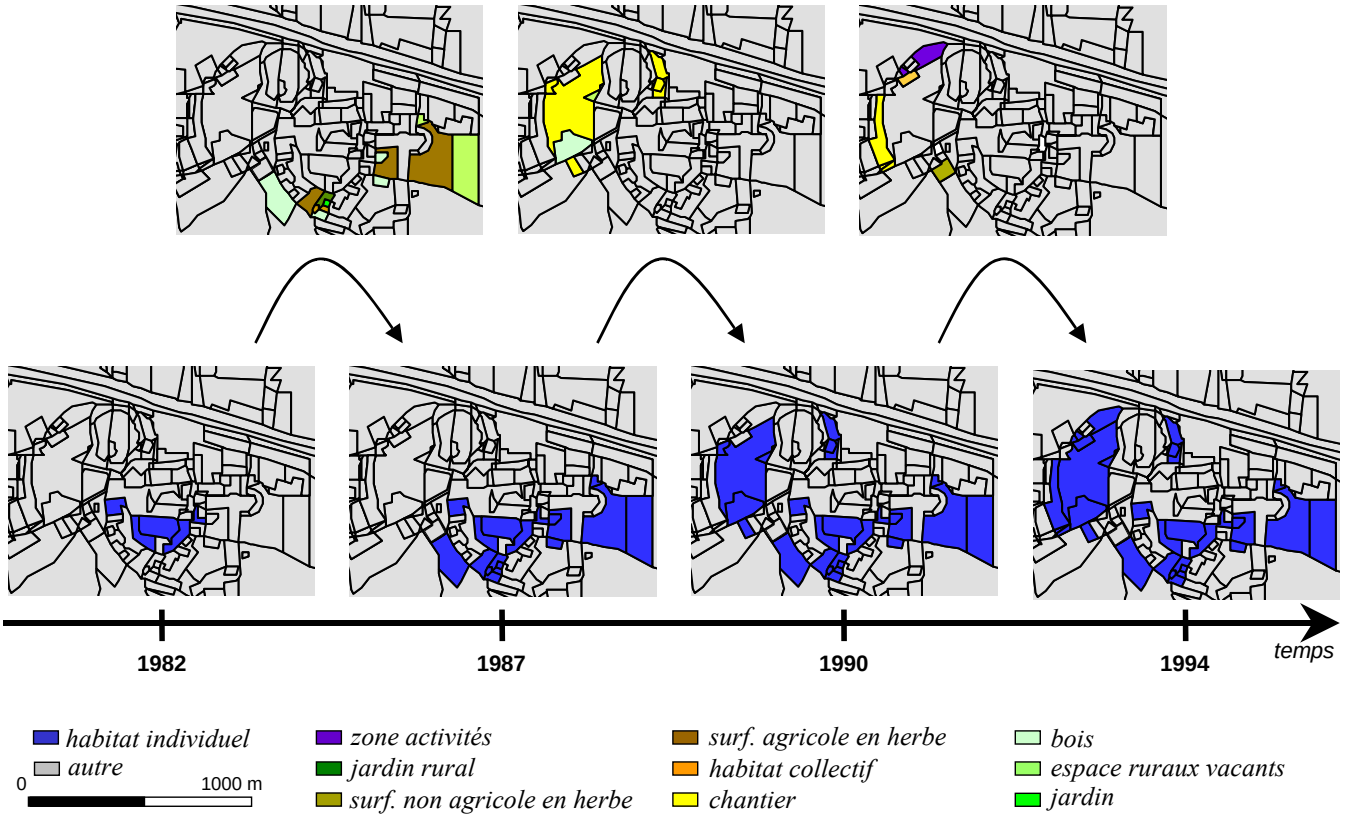


Figure A.2.9. *Complémentarité des descriptions par états successifs et par suite de mutations pour l'observation d'un phénomène*

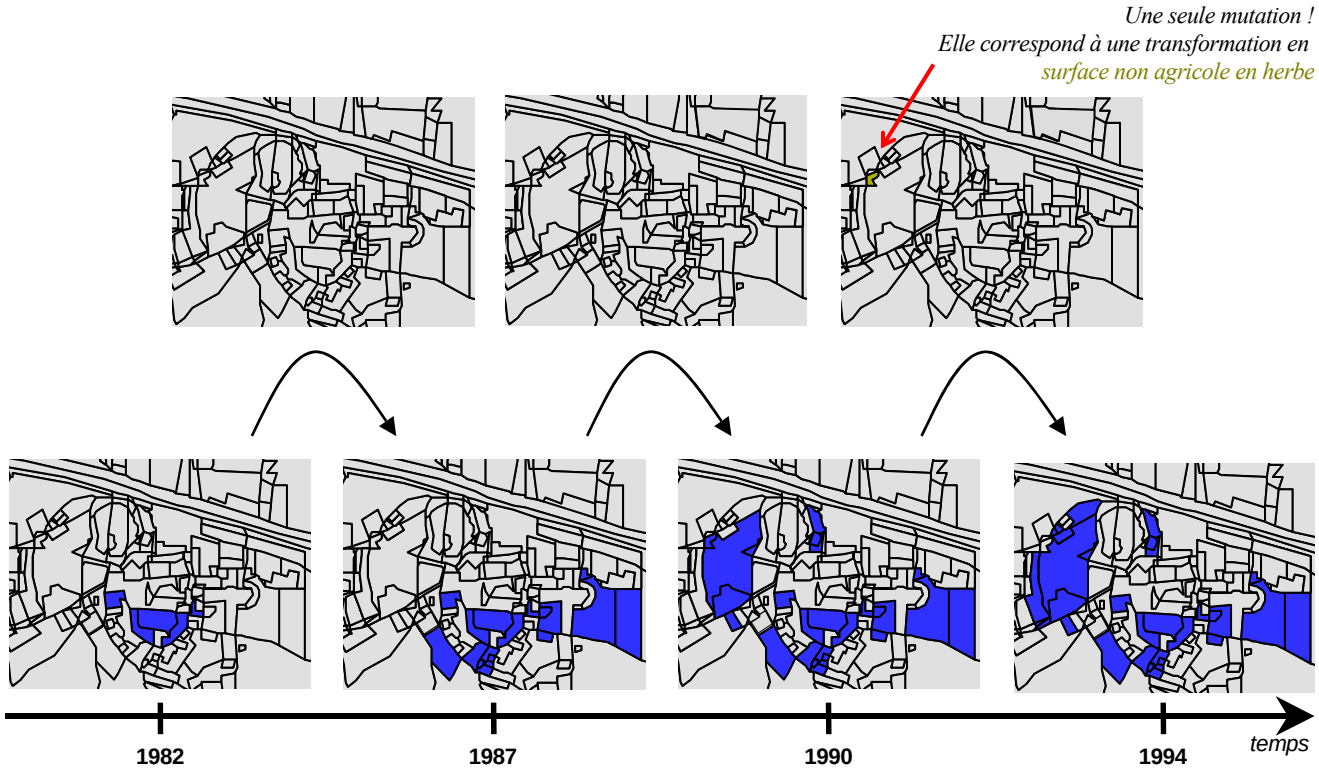


Figure A.2.10. *Description de l'habitat individuel à l'aide des mutations de type destruction*

La figure A.2.9 présente les mutations de type *création*. On peut de même décrire les mutations de type *destruction : « Quelles sont les zones d'habitat individuel qui changent d'occupation des sols entre deux dates? »* (fig. A.2.10). Sur la portion de territoire étudiée, on constate que le phénomène est pérenne : les zones d'habitat individuel une fois installées demeurent. En effet, il n'a y a aucune mutation sur les deux premières périodes et une seule 'zone' change au cours de la troisième période. Ainsi, le développement constaté à partir des mutations de type *création* n'est atténué par aucun mouvement inverse du phénomène qui serait dû à des disparitions, ce qui explique le solde positif constaté sur les états.

Dans l'exemple de la figure A.2.11, on ne s'intéresse plus à l'évolution d'un type particulier d'occupation du sol ; on considère l'ensemble des évolutions, tous types confondus. La partie inférieure fait l'inventaire de différents types, tandis que la partie supérieure montre où se trouvent les mutations. Cette dernière description n'est pas détaillée pour des raisons cartographiques (une légende spécifiant les n*(n-1) (soit 83*82= 6806) cas possibles de mutations n'est pas envisageable). Cet exemple montre que, lorsque la description par mutation est réduite au minimum *(« quels sont les objets qui ont changé ? »*), il devient indispensable de posséder une description détaillée des états pour avoir une représentation complète du phénomène. A l'inverse, l'exemple précédent montrait qu'une représentation par mutation peut compléter une description par état sommaire (dans la figure A.2.9, la représentation par état se contentait de recenser les 'zones pavillonnaires'). La complémentarité résulte de leur dualité : on peut venir enrichir l'une en détaillant l'autre.

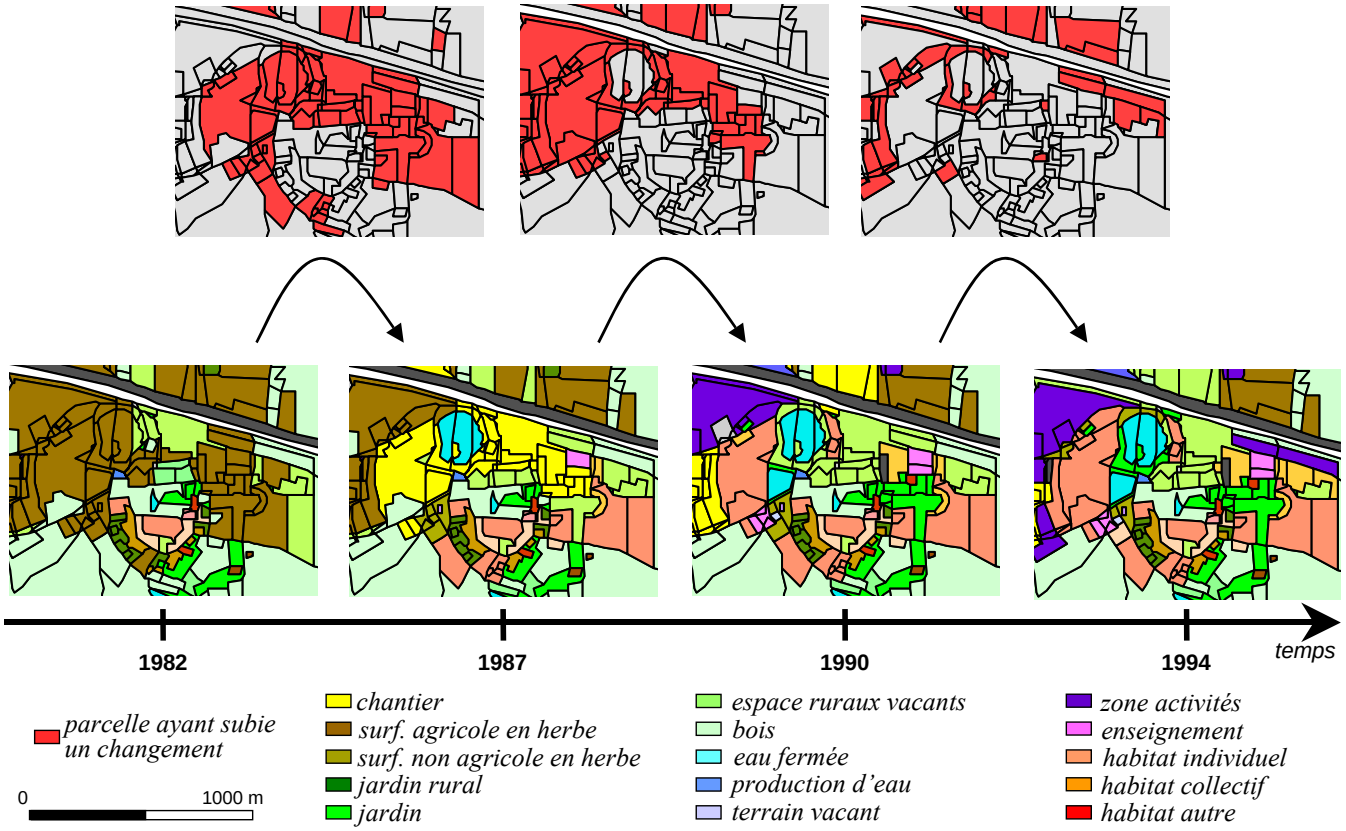


Figure A.2.11. *Description des changements d'occupations des sols sur un territoire*

Dans ces travaux, nous proposons de développer :
une *description* discrète des phénomènes *sous forme d'états* successifs,
qui permette un *suivi automatique* des *mutations* dans le SIG,
pour pouvoir analyser les évolutions.

A.2.1.4. Observation de la dimension spatiale et de la dimension temporelle

A.2.1.4.1. Caractérisation des observations

Une représentation discrète peut décrire des phénomènes continus. Par exemple, le suivi d'une inondation peut se faire à intervalles réguliers à partir de relevés photographiques ou sur le terrain. En effectuant des observations rapprochées, on peut même rendre l'impression de continuité. Nous avons évoqué le rendu de la continuité temporelle avec l'exemple des films cinématographiques, où l'intervalle de temps séparant deux images doit être inférieur à la capacité de traitement de celles-ci. De même pour rendre l'impression de continuité spatiale dans une photographie numérique, il suffit que les pixels soient suffisamment petits pour ne pas être distingués. Le terme de *granularité* permet d'évoquer cette caractéristique des observations. Il désigne l'intervalle minimal entre deux observations et permet alors de parler de *granularité temporelle* et de *granularité spatiale*.

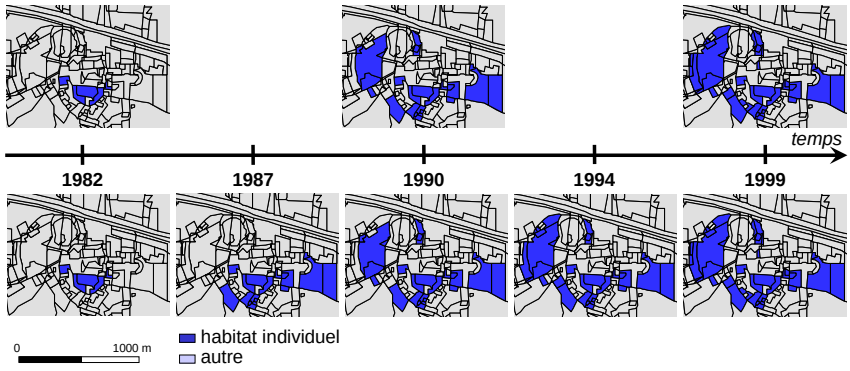


Figure A.2.12. *Une augmentation de granularité temporelle dans la description de l'occupation des sols*

Dans une approche générale :

– La *granularité temporelle* désigne le *pas temporel*, soit l'intervalle de temps minimal séparant deux observations. La figure A.2.12 par exemple présente les données du MOS avec deux granularités temporelles différentes : tous les 8-9 ans en haut, tous les 4-5 ans en bas.

– La *granularité spatiale* caractérise les objets élémentaires observés (par exemple le *pas* ou taille du pixel en mode raster).

Toutefois d'autres types d'intervalles (spatiaux ou temporels) existent, dont la distance minimale participe également à la caractérisation d'une observation. Ainsi, le terme *granularité* prend parfois des sens différents, lorsqu'il ne cède pas la place à d'autres notions (cf. annexe 1).

Granularité temporelle

[Frank, 1998, p.53] donne par exemple une autre illustration de la *granularité temporelle* (cf. figure A.2.13), qu'il intitule *« time with differing granularity »*. Dans cette figure, la représentation de la granularité journalière se présente sous forme d'écart entre deux ponctuels comme dans la représentation de la figure A.2.12. En revanche la représentation des granularités hebdomadaire et annuelle introduit un nouvel élément; la *durée d'observation*. Jusqu'à présent nous avons présenté l'observation comme ponctuelle, mais elle peut posséder une « épaisseur », sorte d'équivalent du temps de pose (ou d'exposition) en photographie, qui se distingue de la période ou durée séparant deux prises de vue lors d'une prise en rafale[5]. Cette différence entre *date* et *durée* de l'observation se retrouve et s'enrichit encore de variantes supplémentaires lorsque l'on s'intéresse à une observation des phénomènes au travers des données géographiques. En effet, les processus dans la production des données (la collecte des données, l'intégration) ne sont pas instantanés. Le producteur cherche à réduire au minimum les délais de production. Souvent incompressibles, ils nécessitent cependant un « intervalle de temps minimum ». La note 6 évoque ces granularités temporelles supplémentaires qui interviennent dans l'observation d'un phénomène à partir de données, en présentant certains aspects spécifiques de la production des données. Elle montre comment la modélisation des phénomènes en données géographiques introduit de fait des décalages temporels. Par exemple, lorsque la date de mise à jour associée à un 'objet' (une donnée) est une année, elle ne renseigne pas sur la durée (intervalle temporel) de validité de la représentation (l'année entière). Elle correspond à une date (un ponctuel

5. Remarque : Dans l'espace temporel, la durée représente l'emprise temporelle des événements, équivalent de l'emprise au sol des objets géographiques, tandis que la date (qui est un « ponctuel » dans l'espace temps) est l'équivalent de la localisation (par exemple par le centroïde) des objets géographiques.

6. Le temps incompressible nécessaire à la collecte des informations de mise à jour soulève des problèmes spécifiques pour les producteurs de données. En particulier, certains événements participant à l'évolution du territoire ont une durée telle qu'ils nécessitent des spécifications propres.
Il s'agit par exemple d'événements dont la durée de vie est inférieure au temps des mises à jour, ils ne sont alors pas pris en compte. Par exemple l'information sur la présence d'un chantier par essence temporaire, peut être devenue obsolète à la sortie de la nouvelle version des données. Cette remarque peut conduire à définir une autre granularité temporelle, conçue comme « la durée minimale que doit durer un fait pour être saisissable ». Cette granularité simple à définir dans le cas de mises à jour régulières (ex.: « l'événement doit durer au moins un cycle de mise à jour »), demande une réflexion thématique plus importante dans le cas d'une *mise à jour dite en continue* (où tout changement significatif est introduit au plus tôt).
Parallèlement, certains événements ont une durée de création (à ne pas confondre à la durée de vie), supérieure au temps de mise à jour. Ils peuvent débuter avant la mise à jour et ne pas avoir abouti à sa fin. C'est le cas par exemple avec la création d'une autoroute qui s'étale sur plusieurs années. Le problème qui se pose alors est de savoir quand introduire l'information dans la base de donnée (ex. : au moment de l'avant projet, quand est connu le fuseau d'implantation, lors de son ouverture au public), sachant qu'au temps de collecte des informations doit s'ajouter un temps d'intégration qui diffère d'autant la mise à disposition de cette information souvent capitale. Une autre granularité temporelle peut alors être ici envisagée : « celle correspondant à la durée de mutation », qui conduit souvent à décomposer un événement en sous événements / étapes dont la durée est plus cohérente avec le processus de mise à jour.
Ces réflexions illustrent le travail technique de modélisation sous-jacent à toute observation de phénomènes géographiques dans un SIG. Elles donnent des exemples des problèmes sur lesquels il s'agit de statuer pour produire des données à étudier.

temporel) d'apparition (dans la base ou sur le terrain). Ainsi, la représentation peut correspondre à un état au début de l'année qui a changé ensuite ou au contraire à un état en fin d'année (voire à un état anticipé sur l'année suivante, cf. le cas de l'autoroute).

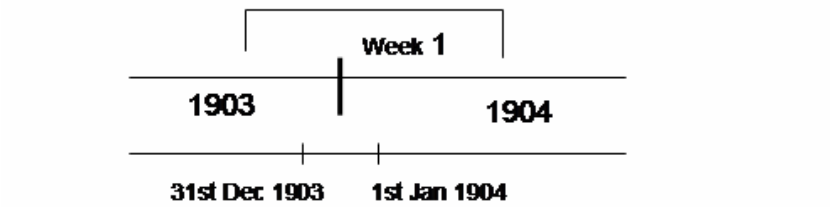


Figure A.2.13. *granularité temporelle*
« Time with differing granularity" : figure extraite de [Frank,1998]

Caractérisation spatiale

L'exemple du pixel en photographie a introduit la *granularité spatiale* comme un terme *technique* dans le contexte d'une représentation *raster*. On retrouve également le terme de *granularité* en représentation *vecteur* ; il se confond parfois avec deux autres caractéristiques spatiales : la *précision* et la *résolution* [Ruas et Bianchin, 2002](cf. annexe AA.I. pour une présentation comparée de ces trois notions). Mais on peut également définir une granularité « plus thématique » avec un *grain* qui ne soit pas une *caractéristique de la représentation*, mais de *ce qui est représenté* [Hernandez, 1995, p.49].

Ainsi, la *granularité spatiale* qui « caractérise l'objet élémentaire observé » peut renvoyer :

– soit à l''objet' en tant que *donnée (représentant)*. La caractérisation de l'observation renvoie alors à des paramètres de qualité des données (« granularité technique »). Elle sert à qualifier la représentation sous forme de données de l'objet réel (ou plutôt de l'*objet nominal* tel que décrit par les spécifications). La granularité s'ajoute alors à la précision et à la résolution pour caractériser la base de données.

– soit à l'objet géographique en tant qu'*entité (représenté)*. La caractérisation sert alors à préciser quelles sont les *entités* élémentaires observées (« granularité thématique ») (ex. : les individus, les ménages, les îlots, par exemple la CNIL restreint de fait la *granularité thématique*)

La « granularité thématique » (en relation avec le *niveau d'observation* ou l'échelle d'observation) spécifie le plus petit élément géographique participant à l'observation du phénomène étudié, *«c'est à dire les objets qui peuvent être considérés comme des atomes relativement au phénomène étudié, qui sont conceptuellement indivisibles »* [Hypergeo.free.fr]. Dans un même champ thématique, suivant la question, un même phénomène peut en effet être considéré à des échelles ou niveaux différents. Ainsi, dans un questionnement relatif à la régularité de la distribution des tailles des villes d'un pays [Batty, 2001] d'une part, et [Allen et Sanglier, 1979], [Bura et al, 1996] d'autre part adoptent deux positions différents : dans le premier cas le niveau élémentaire est celui de l'individu, dans le second, c'est celui des villes et des villages. La figure A.2.14 est une autre illustration de la variété d'échelles d'observation (à la fois spatiales et temporelles). La détermination du niveau d'observation est une étape thématique préalable à l'application de nos recherches qui est importante. En effet, elle spécifie les entités géographiques à analyser et pose ainsi des contraintes sur les données qu'il faut posséder pour réaliser l'observation.

Le thème de la granularité, au delà des décalages temporels évoqués dans le paragraphe sur la granularité temporelle est symptomatique d'un *décalage* plus important entre *entités* et *données*. Dans ce « gap », le vocabulaire change de signification[7]. Les problématiques elles-même évoluent (nous le montrerons à propos de la problématique du temps dans les SIG –chapitre A.3), faisant apparaître deux pôles dans l'observation :

- un pôle « thématique », où l'observation porte sur des entités pour étudier un phénomène
- un pôle « technique », où l'observation porte sur des données et tente d'en extraire des informations utiles à l'étude du phénomène.

7. La « granularité thématique » devient « technique », l'objet - entité devient 'objet'- donnée, la représentation comme action (la représentation du phénomène) devient la représentation comme résultat (les données comme représentation), etc.

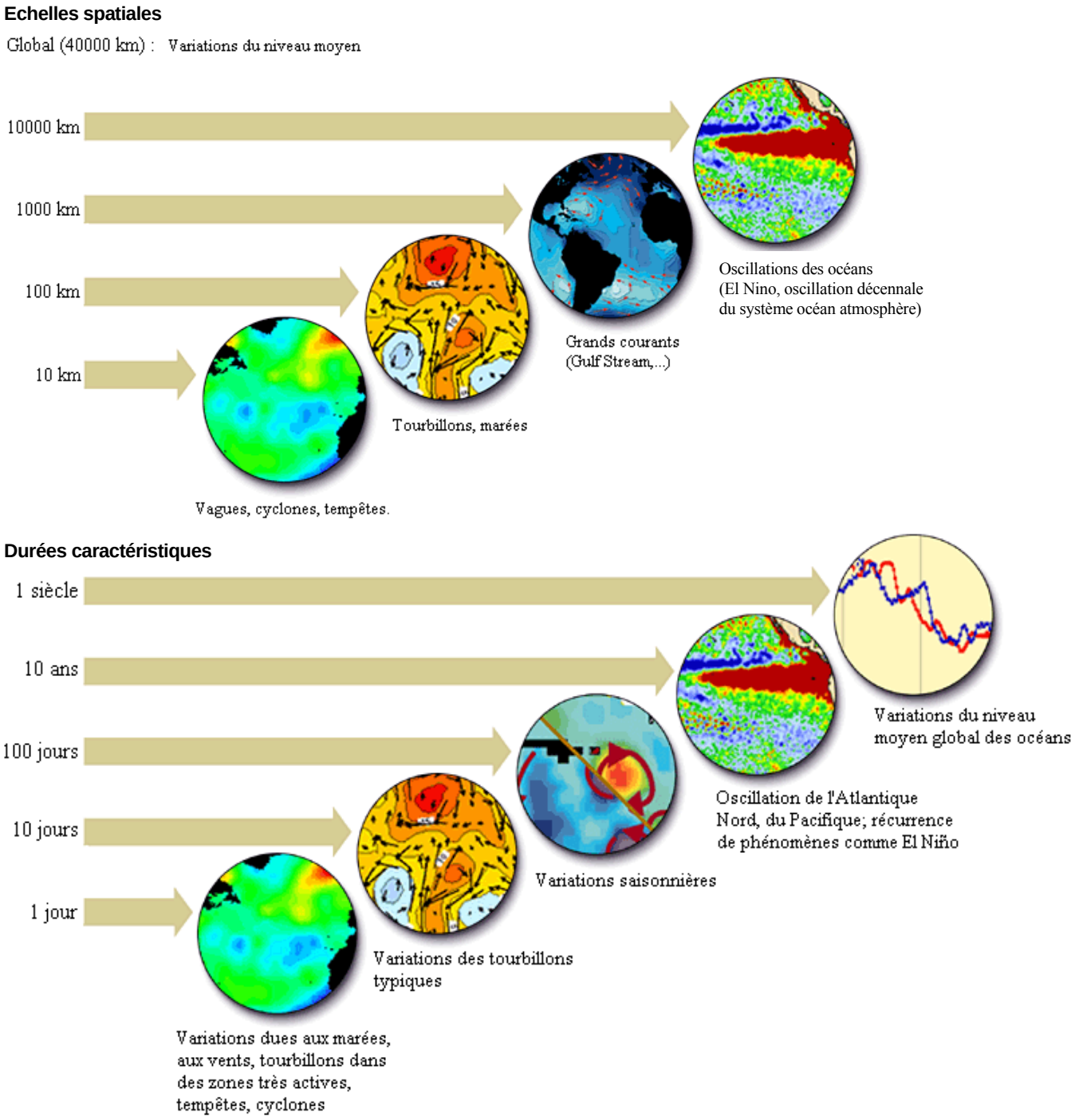


Figure A.2.14. *Différentes échelles d'observation spatiales et temporelles pour étudier les variations altimétriques des océans*

« L'océan sous toutes ses facettes » [Extrait de www.jason.oceanobs.com/html/kiosque/dossiers/alti/facettes_fr.html]
« L'océan, toujours en mouvement, varie selon des rythmes multiples, rendant ainsi nécessaire à sa compréhension une observation en continu. De plus, les zones affectées sont de dimensions variables. L'altimétrie est le reflet de nombre de ces phénomènes océaniques, en mesurant régulièrement et sur tout le globe la hauteur des mers à quelques centimètres près ».

A.2.1.4.2. Observations relatives

Le temps et l'espace jouent des rôles différents selon le type d'observation et de description des phénomènes. Dans l'approche continue, temps et espace semblent tenir des rôles symétriques et interdépendants. Dans l'approche discrète telle que nous l'avons présentée, le temps sert de jalon pour situer les états ou leurs mutations, il sert de support. Dans la pratique, beaucoup d'applications privilégient l'un des aspects (temporel ou spatial) au détriment de l'autre. Plusieurs raisons peuvent expliquer cela :

− les notions de temps et d'espace ont longtemps été considérées indépendamment[8]. Aujourd'hui, on a pris conscience de l'intrication des deux dimensions. Cependant, notre structure mentale elle-même effectuerait des traitements différenciés[9].

− pour effectuer les analyses, l'utilisateur est tributaire des outils, méthodes et/ou modèles qu'il emploie. Généralement, il y est difficile de représenter simultanément le temps *et* l'espace. Les outils, modèles ou méthodes utilisés, imposent d'ailleurs souvent des contraintes en fonction de leurs (in)capacités à considérer conjointement les deux composantes. C'est en particulier le cas des SIG.

− les phénomènes sont souvent des objets complexes. Les décomposer en éléments plus simples, est une démarche scientifique classique, dite réductionniste[10], qui vise à réduire cette complexité. Ainsi, régulièrement on tend à négliger l'une des deux dimensions pour étudier l'autre. Soit on fixe la dimension temporelle pour mettre l'accent sur la dimension spatiale (ex. : localisation des espaces agricoles les plus sensibles à la périurbanisation), soit on fige la dimension spatiale et l'on observe la dimension temporelle (ex. : vitesse d'extension de la zone périurbaine). Dans certains cas, la question même que l'on se pose implique de privilégier l'une des deux dimensions. Par exemple :

- l'étude de la dimension spatiale du phénomène en fonction du temps, convient bien aux problématiques du type : « définir quelle est la *forme* de l'emprise du phénomène à une date t donnée »,
- l'étude de la dimension temporelle du phénomène en fonction de l'information spatiale, convient bien aux problématiques du type : «définir à quelle *date* le phénomène a telle superficie ».

En illustrant nos propos à l'aide de *faits spatiaux* au détriment des *faits temporels*, en utilisant le temps comme jalon pour situer ces états, nous avons déjà pris position en faveur de la dimension spatiale. La contrainte liée aux outils est la raison principale : nous intégrons le fait que l'observation doit s'effectuer « *avec un SIG »* et nous anticipons sur le fait que les SIG privilégient l'espace par rapport au temps [11]. Cependant, *privilégier* l'un ne signifie *pas renoncer à* l'autre. Dans ces travaux, nous nous adaptons à l'outil SIG, pour mieux l'adapter à l'observation des phénomènes[12].

En particulier, nous constatons que la « granularité thématique » établit un lien entre la dimension spatiale et la dimension temporelle : dans une représentation discrète d'un phénomène, le pas d'observation ne peut en effet pas être choisi indépendamment du niveau d'observation. En reprenant l'analogie avec le film, un pas de temps petit donne à voir un certain immobilisme. A l'opposé, avec un pas très important, la représentation de deux faits successifs donne l'impression d'un saut, d'une discontinuité dans le mouvement, d'une rupture, comme si l'objet d'observation n'était plus le même. Il importe de disposer d'un pas de temps suffisamment fin pour saisir des changements qui ne soient pas radicaux et suffisamment important pour que l'observation soit

8. Ce n'est qu'à la Renaissance, avec la révolution copernicienne, que l'idée d'étudier conjointement l'espace et le temps a été généralement admise: *"(with the Copernican revolution), the notion that space and time could be studied together as related aspects of an unified Cosmos became generally accepted"* [Peuquet, 2002, p.17]. Jusqu'alors ils faisaient l'objet de regards distincts et indépendants, s'appuyant sur des représentations disjointes.

9. Des études en neurologie (citées par [Peuquet, 2002, p.197]) semblent indiquer que de l'organisation fonctionnelle de notre cerveau résulterait une séparation des deux perceptions (temporelle et spatiale). En effet, une division du travail s'effectue entre les lobes temporal et frontal : là où l'un encode les propriétés des objets comme la couleur, la forme, etc. l'autre code les propriétés spatiales comme la localisation et la distance… Dans le cerveau, la dimension temporelle suivrait des processus de traitement différents de ceux de la dimension spatiale.

10. *« On appelle réductionnisme, toute démarche philosophique qui consiste à considérer que l'on a compris la totalité de l'objet à partir du moment où il a été réduit à ses éléments les plus fondamentaux. Un réductionniste accomplit une analyse (c'est à dire une démarche qui consiste à diviser un objet en une multitude de parties étudiées séparément) et oublie que la synthèse (qui sert à combiner les éléments) reste à réaliser pour comprendre le fonctionnement des choses »* [Serres et Farouki, 1998].

11. Le chapitre suivant montre cet avantage donné à l'espace par rapport au temps, mais le chapitre A.1 le laissait déjà deviner : il existe une composante géométrique et topologique, mais dans les SIG aucune composante temporelle n'est explicitement évoquée.

12. Ainsi, nous proposons de partir de représentations par états successifs, sans mutation, car ce sont les plus simples à gérer dans les SIG et de fait les plus classiques, mais nous développons une méthode qui permette d'accéder facilement aux changements pour pouvoir observer des évolutions.

intéressante. Dans un contexte où la dimension temporelle est contrainte, cette condition peut alors s'interpréter dans un sens inverse : « à pas d'observation fixé, il faut disposer d'un niveau d'observation suffisamment fin pour saisir des changements… ». Ce qui sera intégré dans le chapitre B.3 sous la forme : « L'introduction d'un niveau d'observation plus fin, peut aider à saisir des évolutions de détails ».

A.2.2. Etude de phénomènes

Tôt dans ce chapitre nous révélons la nature et la forme de l'observation développée dans ces travaux : *une observation à partir d'états, qui permette un suivi automatique des mutations pour décrire des phénomènes.* Dans ce paragraphe, nous situons cette observation parmi les études (temporelles et spatiales) de phénomènes, pour montrer comment elle peut participer aux analyses de ces phénomènes.

Dans un premier temps, nous introduisons la notion de *système de référence* pour évoquer les différents regards portés sur les phénomènes. Elle nous sert également à introduire une typologie des études de phénomènes, proche des sciences physiques, centrée autour du statut du mouvement. Nous nous appuyons ensuite sur cette typologie pour préciser la méthode d'observation proposée.

Dans un second temps, nous justifions a posteriori le modèle temporel implicitement retenu pour les observations (la ligne orientée). Puis nous nous situons parmi les études temporelles en confortant la proposition d'une observation descriptive des évolutions passées.

A.2.2.1. Etudes de la dimension spatiale

L'observation d'un phénomène s'inscrit généralement dans le cadre d'une étude plus large (cf. paragraphe A.1.3.2). Avant de proposer une méthode d'observation pour un phénomène, il importe de savoir ce que l'on veut étudier, de connaître le *système de référence* de l'observateur. En particulier, pour un phénomène considéré comme un *fait en mouvement*, il s'agit de préciser :

- le fait observé. Nous avons déjà vu qu'il fallait préciser le *niveau d'observation*, mais il faut également spécifier les caractéristiques observées, (par exemple, l'observation du phénomène d'urbanisation peut être centrée sur les individus, les activités, les constructions, etc.)
- le statut du mouvement. Le mouvement en tant que tel, fait-il partie de l'observation ? Veut-on le décrire ? Veut-on l'expliquer ?

A.2.2.1.1. Systèmes de référence pour l'étude de phénomènes

Le *système de référence* de l'observateur peut être compris dans un sens *technique* lorsqu'il s'agit des données (le système de référence comme système géodésique). En effet, pour localiser les 'objets', il est nécessaire de définir un système de coordonnées (cf. A.1.2.4) Ce système varie alors suivant le lieu du globe (où se trouve le phénomène, où se trouve l'observateur) (fig.A.2.15) et suivant l'étendue du phénomène.

Le *système de référence* peut être compris comme *thématique*, lorsqu'il s'agit de l'observation des entités géographiques (le système de référence comme point de vue) ; des chercheurs issus de disciplines différentes étudient des aspects divers d'un même phénomène. Certains séminaires scientifiques illustrent la complémentarité des approches différentes issues de disciplines variées (ex. : sur le risque, sur le développement durable).

Figure A.2.15. *Système de référence de l'observateur (extrait de « En haut et en bas » d'Escher)*

La spécification des informations géographiques utilisées pour observer un phénomène dans un SIG implique de préciser l'ensemble des systèmes de référence : les systèmes de référence techniques (car les informations doivent être représentées en *données*) et thématiques (car les informations modélisent les *entités*). Or, si les systèmes de coordonnées sont sans ambiguïté (la géodésie les détermine entièrement) et exclusifs (dans un SIG, des données issues de systèmes différents ne peuvent être observées conjointement sans changement de coordonnées), les systèmes de références thématiques sont plus ouverts, en raison de la *complexité* des phénomènes. Ainsi nous constatons que même si l'observation d'un phénomène se concentre sur un aspect particulier, il existe souvent des relations avec des thèmes connexes. Il est alors possible de compléter les informations avec des éléments issus de systèmes de références thématiques différents sans sortir du système initial. Ainsi, dans le phénomène d'urbanisation pavillonnaire, on peut s'intéresser au bâti (à sa densité, sa forme, etc.), mais on peut aussi introduire le réseau routier (sa structure, son trafic, etc.), le parcellaire (sa forme, son statut juridique, etc.), la population (ses attributs, sa structure), etc., car ces caractéristiques sont également susceptibles de participer à l'évolution du phénomène. D'un point de vue pratique, les SIG n'imposent apparemment pas de contrainte. En effet, ils permettent de présenter toutes les données dès lors qu'elles sont localisées dans le même système de coordonnées[13]. Le thématicien peut donc sélectionner seul, les variables qui seront observées et celles qui feront partie des « choses restant égales par ailleurs ».

13. Dans le chapitre B.2, nous montrerons que les SIG peuvent présenter toutes les données lorsqu'elles sont dans le même système de coordonnées, mais que cela n'implique pas que la présentation soit pertinente. En effet, il faut également que les échelles de représentations soient compatibles (des données au 1/100ième ne peuvent s'afficher avec des données au 1/1 000 000) ; C'est le *problème du multi échelle en cartographie numérique* (qui ne se confond pas avec le problème de l'échelle d'observation en géographie). Plus précisément, il importe que les données soient cohérentes entre elles, qu'elles se rapportent aux mêmes données support. Dans le chapitre B.3, nous introduirons la notion de *système de données*, pour spécifier que des données sont dans un même système de référence géodésique ET qu'elles sont géo-métriquement cohérentes les unes avec les autres.

Figure A.2.16. *Les systèmes de références comme métaphore pour évoquer les différents façons d'observer le mouvement dans une approche semblable à celle des sciences physiques*

A.2.2.1.2. Différents regards sur les phénomènes, différents types d'étude

Il peut exister plusieurs façons thématiques d'aborder un phénomène et il y a plusieurs manières de l'étudier en tant que « fait en mouvement » (fig. A.2.16). Il s'agit ici de préciser si l'observation se concentre sur les faits pour décrire des états (approche statique) ou si elle s'intéresse au mouvement pour le représenter, l'expliquer ou encore le simuler. Ainsi, comme l'évoque [Brunet et al., 1992], le regard sur un phénomène et son étude prennent des formes différentes si l'on cherche à identifier des formes ou structures comme Kant ou si l'on tente de découvrir des lois comme Lautréamont[14]. Pour [Brunet et al., 1992], l'objet de la géographie englobe l'ensemble de ces études : *« La géographe étudie des phénomènes, des systèmes et des processus »*. En nous inspirant des sciences physiques et des différentes approches du mouvement, nous proposons avec {Durand-Dastès, 2005} de distinguer trois types d'étude de phénomènes en fonction de l'interprétation de la dimension temporelle :

- l'**étude statique** comparative reposant sur des instantanés,
- l'**étude cinématique**, qui décrive le mouvement,
- l'**étude dynamique,** qui tente d'expliquer le mouvement.

14. *« Le phénomène passe, je cherche les lois »*, Lautréamont, poésie II.

L'étude statique

Le caractère *statique* renvoie à ce qui est fixé, à ce qui n'évolue pas. L'étude statique s'attache à obtenir des descriptions d'états, sortes de clichés ou *snapshots*[15]. Par essence descriptive, elle décompose les faits en ses éléments constituants et tente d'en donner une représentation structurelle. Elle s'associe à l'approche discrète des phénomènes qu'elle présente sous la forme d'une succession d'états. Elle laisse aux chercheurs le soin d'effectuer les comparaisons pour identifier les mutations et suivre des évolutions. Car l'évolution qui implique un mouvement n'est pas du domaine statique ; elle fait partie d'un autre registre d'études.

15. Le terme de *snapshot* repris au domaine de la photographie a parfois pour équivalent en français : « instantanée ». Il est intéressant de noté ici, que la représentation d'un état dans un base de donnée, n'est pas « exactement » un instantané. En raison de délais de production et d'intégration des données, entre autre. Ainsi l'utilisateur tend à utiliser la représentation d'un état sur une certaine durée, fonction généralement de la mise à jour suivante. On pourrait à l'image de la « persistance rétinienne » qui intervient dans le principe du film, parler ici de « persistance opérationnelle ». Dans le chapitre B.3, nous utiliserons cette remarque pour améliorer les observations temporelles.

L'étude cinématique

La *cinématique* est la partie de la mécanique qui étudie le mouvement indépendamment des forces qui le produisent. Une étude cinématique s'intéresse donc au phénomène comme mouvement et s'attache à le décrire. Elle étudie comment le fait évolue dans l'espace et/ou le temps.

Les études cinématiques peuvent être réalisées dans le cadre d'une approche discrète. Elles introduisent alors les mutations entre états de faits. En fonction de l'objet étudié, les relations entre les mutations et les états sont plus ou moins développées. Ainsi, la description des faits peut n'être que secondaire (connaissance de toutes les mutations et uniquement description du fait initial pour reconstituer la trajectoire). C'est le cas des analyses centrées sur les changements eux-mêmes (ex. : leur rythme, leur forme). En revanche, si l'analyse porte sur les faits en mouvement, mutations et états sont décrits de façon plus coordonnée. On peut par exemple adopter cette approche pour étudier le développement d'une zone urbaine et observer sa forme à travers les états et ses transformations à l'aide des mutations (cf. A.2.9 et A.2.10). Cela peut alors conduire à étudier la succession des états et des mutations, pour tenter de trouver des règles de transformations, pour débuter des études plus dynamiques.

Les études cinématiques peuvent aussi être réalisées dans le cadre d'une approche continue. Le mouvement est alors décrit sous forme d'une fonction (ex. : phénomène de diffusion en fonction du temps et de la distance) et l'étude de sa dérivé, si elle existe, donne des informations complémentaires sur ce mouvement (ex. : vitesse, croissance/décroissance, périodicité) (cf. figure A.2.2).

La description dynamique

La *dynamique* est la branche de la mécanique qui étudie le mouvement d'un mobile considéré dans ses rapports avec les forces qui en sont la cause[16]. Elle s'intéresse à *« l'ensemble des forces en interaction et en opposition dans un phénomène, une structure »* [Petit Robert, 2001]. Les études dynamiques de phénomènes qui peuvent être réalisées dans le cadre d'une approche discrète ou continue, s'attachent à décrire les éléments en jeu dans le phénomène, ceux qui l'expliquent en tant que mouvement (analyses descriptives). Elles tendent à comprendre la genèse et le processus du phénomène (analyses explicatives). Certains travaux ont pour objectif de reproduire « un fait stylisé » [Barros, 2003], d'autres de reproduire un phénomène observé [Holm et al, 2004], [Moeckel et al., 2002]. Les premiers cherchent à identifier un ensemble de règles simples qui aboutissent à des structures observables et durables (ex. : le modèle de [Shelling, 1971]), les autres tentent de formaliser l'ensemble des règles et d'estimer les valeurs des paramètres qui permettent de reproduire les dynamiques observées [Sanders, 2006]. Dans ce dernier cas, la confrontation avec les observations empiriques devient une étape

16. Il est intéressant de remarquer ici, que la dynamique prend parfois en géographie et en sociologie, une signification qui diffère un peu de celle donnée en sciences physiques : elle désigne le domaine étudiant spécifiquement les faits en évolution. Les études dynamiques s'identifient alors aux études de phénomènes évolutifs sans préciser si l'on s'intéresse au mouvement ou à ses causes. Nous avons choisi de ne pas adopter cette définition pour conserver dans toute la suite, le sens attribué au terme d'étude dynamique proposé ici. Il importe en effet pour nous de pouvoir faire une distinction entre les trois types d'analyse de phénomènes évolutifs.

essentielle. Ces études reposent sur une bonne connaissance des phénomènes portant à la fois sur leur forme (connaissance résultant d'analyses cinématiques) et sur leurs causes (connaissance sur des éléments connexes). C'est pourquoi, elles requièrent généralement une expertise thématique spécifique.

Parmi l'ensemble des approches dynamiques, certaines s'appuient sur les récents développements en informatique pour effectuer des simulations ([Batty et al, 1997], [O'Sullivan et Torren, 2000], [Benenson et al, 2002] cités in [Benenson et Torrens, 2004-II]). [Benenson et Torrens, 2004-I] emploient le terme de *géosimulation* pour désigner un nouvel ensemble de concepts, de méthodes et techniques, qui combinent les apports de la théorie des systèmes complexes, des sciences de l'information géographique et de la programmation orientée objet. Pluridisciplinaires, ces modèles de géosimulation empruntent à plusieurs domaines scientifiques techniques (géomatique, intelligence artificielle, techniques de visualisation 3D, modèles mutli-agents etc.) et thématiques (modèles de transports ou d'occupation des sols, modèles de diffusion spatiale, etc.). Ils utilisent souvent les automates cellulaires qui permettent de *« formaliser l'espace sous la forme d'une grille régulière de cellules, chacune d'entre elles étant caractérisée par un état, état évoluant en fonction des caractéristiques des cellules voisines suivant des règles de transition »* [Hypergéo.free.fr]. La recherche urbaine est l'un des domaines où ces modèles se sont le plus développés. Partageant avec la *microsimulation* une approche individu-centrée, la géosimulation :

- utilise des modèles qui fonctionnent sur des objets élémentaires discrets, à une échelle « microscopique » (ex. : maison, parcelles) : *« geosimulation considers urban systems as collectives of spatially non-modifiable elements, or objects, at « atomic » resolutions : individual people, households, vehicles, buildings, land parcels, etc. »* [Benenson et Torrens, 2004-II]. Elle ne travaille pas à partir d'unités spatiales *modifiables* (c'est à dire *formées d'unités élémentaires d'une partition de l'espace* [Cheylan et al, 1994]),

- intègre des règles d'interaction entre les objets élémentaires, sans se limiter à un nombre limité d'assomptions. Elle considère l'émergence de structures durables à d'autres niveaux d'organisation et place les mécanismes « bottom-up » à la base de la compréhension des structures spatiales.

- là où traditionnellement les modèles urbains incluent des visions spatiales simples issues de *snapshots*, avec un nombre restreint de temps de référence, les modèles de type géosimulation travaillent à des échelles de temps proches du temps réel pour construire des simulations dynamiques. Pour cela, les temps d'intégration des informations de mises à jour sont réduits (par exemple le processus de production des données individualise l'implémentation des changements au plus près de leur survenue, au lieu de planifier des mises à jour périodiques de la totalité de la base). La granularité temporelle des observations est alors plus fine et plus proche de la « réalité », permettant en simulation d'explorer une plus large gamme d'hypothèses.

- enfin, le but des modèles de géosimulation diffère de celui des modèles de simulation traditionnels qui visent la prédiction. En effet, leur objet est moins l'anticipation que la prospective. Ces modèles servent avant tout à tester des scénarios et sont plutôt conçus comme des *« tools to think with »* (outils d'aide à la réflexion) avec pour objectif de tester des hypothèses relatives aux observables à d'autres niveaux d'organisation géographique.

Dans le chapitre A.1, p.39, nous nous sommes inscrits dans la perspective de SIG-*outils d'aide à la décision*. Ces outils nous les concevons à la fois dans une perspective d'outils de « fouille de données » (c'est à dire des outils capables d'extraire des nouvelles informations) et dans l'idée *d'outils d'aide à la réflexion* (pour valider de nouvelles informations). Pour élaborer, tester et valider de nouvelles informations, une première étape consiste déjà à disposer d'informations « dynamiques » (« dynamique » étant pris ici dans le sens courant de « relatif au mouvement, sous-entendu sa forme comme ses causes »). Cette thèse vise à faire évoluer les observations souvent statiques dans les SIG-observatoire vers des analyses plus « dynamiques ». La passage par l'étape *cinématique* nous paraît être une étape intermédiaire (d'où l'importance d'un vocabulaire différent pour désigner le caractère cinématique du dynamique) : il s'agit déjà d'intégrer le mouvement et les évolutions dans les modèles utilisés en standard par les SIG. Dans un second temps, on pourra étudier comment introduire les causes de ce mouvement et de ces évolutions afin d'effectuer les développements que l'absence de modélisation des événements dans les SIG limite aujourd'hui.

De l'approche dynamique en géosimulation, nous retenons :

- l'idée d'une analyse à un niveau élémentaire (micro) pour décrire un niveau supérieur,
- l'intérêt pour des unités géographiques non modifiables,
- l'intérêt pour des observations temporelles moins dépendantes de dates prédéfinies (celles des *snapshots*),

Notre choix de travailler en mode vecteur pour tirer profit des la « modélisation orientée objet » des outils SIG, est compatible avec l'approche de la géosimulation : *« geosimulation as object-based modelling»* [Benenson et Torrens, 2004-II]. Néanmoins, il revient à travailler sur des « cellules » qui sont des objets géographiques élémentaires et non des mailles d'une grille régulière. Les informations fournies par la méthode d'observation proposée ne seront donc pas celles attendues par les automates cellulaires (outils utilisés couramment en géosimultation qui travaillent sur des cellules géométriquement régulières). Toutefois, dans la perspective de travaux en géosimulation, elles conviennent a priori à d'autres modèles : multi-agents.

A.2.2.2. Etude de la dimension temporelle : les phénomènes évolutifs

En introduction de ce chapitre, nous avons évoqué le rôle de la dimension spatiale et de la dimension temporelle dans les phénomènes géographiques. Nous avons cependant ensuite rapidement privilégié la dimension spatiale pour préciser la nature des observations proposées dans ces recherches. Dans ce paragraphe, nous revenons à la dimension temporelle pour présenter certains aspects particuliers liés à l'étude des phénomènes qualifiés d'« évolutifs ». Dans l'observation des phénomènes évolutifs, le temps n'est plus seulement une dimension pour situer les faits ; au travers de l'étude des évolutions, le temps devient explicitement une dimension d'analyse.

A.2.2.2.1. Observation de la dimension temporelle

Selon le dictionnaire [Petit Robert, 2001], une évolution est *« formée de changements successifs insensibles »*, ce qui semble nous placer dans un contexte continu. Dans une approche continue physique, pour observer les

évolutions d'un phénomène, on s'intéresse à la dérivée de la fonction qui décrit le mouvement étudié. On étudie ses caractéristiques, par exemple on cherche les points d'inflexions, les intervalles de croissance / décroissance, de constance, on étudie la continuité du mouvement.

En tant que *« transformation graduelle assez lente »*, l'évolution peut cependant s'envisager également dans une approche discrète. L'étude d'un phénomène évolutif porte sur sa forme temporelle. Dans une approche discrète, on s'intéresse alors aux évènements, identifiés à des faits temporels. Par exemple dans une analyse statique, on observe leur localisation, leur type, leur répartition dans le temps. A l'aide d'une analyse cinématique, on regarde leur enchaînement, on reconstitue le mouvement du phénomène dans le temps. Plus généralement, on tente de découvrir un sens au mouvement, le mot *sens* signifiant ici une direction : «le phénomène va-t-il en s'accroissant, en diminuant dans le temps, évolue t'il de façon chaotique, est-il périodique ? ». On regarde les variations de quantité, de direction, de rythme, etc. Dans une analyse plus dynamique, on s'intéresse au *sens* des phénomènes évolutifs, où le sens est compris comme une *signification*. Ainsi, certains considèrent que chaque position spatiale, en s'intégrant dans une suite de transformations, participe non seulement d'un mouvement dans l'espace, mais également d'un mouvement plus global, que l'on pourrait appeler l'histoire du phénomène ou son processus [Kelmelis, 1998]. Choisissant d'introduire les faits spatiaux, sans se limiter à l'étude de la forme temporelle du phénomène, ils tentent alors à l'aide d'analyses cinématiques et dynamiques d'étudier son histoire, voire son devenir.

L'étude des phénomènes évolutifs s'intéresse aux changements : à leur localisation dans le temps, à leur nature et à leur origine. Ces informations ne sont pas toutes disponibles dans un SIG. Lorsque la description est sous forme d'état successifs, les changements ne sont pas directement accessibles. Seule l'intégration des mutations permet de les expliciter et leur localisation temporelle est alors souvent imprécise[17]. La description des caractéristiques des changements dépend de la modélisation des mutations. Quant à la description de la forme et de l'origine de ces changements, elle est généralement inexistante faute de modèle pour les intégrer. (Ainsi faute d'informations spécifiques sur les mutations, rien ne permet par exemple avec les données du recensement de déterminer si un individu qui habite à Paris en 1999 et en 2004 n'a pas déménagé entre temps et pourquoi il est à Paris en 2004). La modélisation dans les SIG contraint ainsi l'observation des phénomènes évolutifs. L'intégration des mutations devient alors un moyen pour introduire des éléments sur les faits temporels dans une représentation par ailleurs presque exclusivement spatiale.

17. Quand rien de particulier n'a été prévu pour spécifier la date précise du changement, sa localisation temporelle est généralement donnée par l'intervalle de temps décrit par les deux états (initial et final) de la mutation.

A.2.2.2.2. Systèmes de référence temporelle

La plupart des analyses de phénomènes évolutifs s'appuient sur la modélisation classique du temps: l'espace temps linéaire à une dimension, orienté, qui mène du passé vers le futur [Peuquet, 2002] (fig. A.2.17) (ex. : [Denis, 1998], [Madoré, 2001], [Bonnefoy et Goiffon, 2003]). On notera que le caractère orienté de ce modèle permet en particulier de répondre au besoin d'ordre implicitement contenu dans la vision des évolutions comme «suite de transformations » ou encore « changements successifs ».

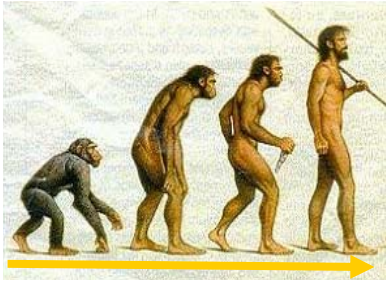


Figure A.2.17. *Le modèle temporel de référence : le modèle linéaire à une dimension, orienté*

Pour ces travaux, nous utilisons également ce modèle pour sa simplicité et sa pertinence pour le plus grand nombre d'utilisateurs. Toutefois, nous n'excluons pas, dans le cadre de développements ultérieurs, de recourir à

des modèles plus élaborés. D'autres modèles existent en effet qui peuvent s'avérer plus adaptés dans certaines applications [Worboys, 1998], [Frank, 1998]. Par exemple, le temps comme boucle rend mieux compte du caractère cyclique de certains phénomènes comme les saisons ou les rythmes quotidiens [Banos et Thevenin, 2003] (fig. A.2.18). Le modèle par branche est également intéressant dans le cas d'études prospectives, car il permet d'envisager simultanément des hypothèses d'évolutions variées.

Figure A.2.18. *Deux modèles conceptuels différents du temps*

[Pastor, 2004] recense ainsi quatre types de représentation du temps :

- la séquence linéaire (passé, présent, futur) ;
- le cycle (représentation des rythmes récurrents avec périodicité régulière, par exemple cycles naturels, biologiques, sociaux...) ;
- un ensemble de trajectoires potentielles et parallèles avec des bifurcations ;
- le temps multidimensionnel pour distinguer le moment où un phénomène survient dans la réalité (temps réel) de ceux où il est observé ou mesuré (temps perçu), renseigné dans la base de données (temps de transaction), représentatif de la situation réelle (temps de validité) ou utilisé pour faire un traitement (temps d'usage) .

La modélisation de la dimension temporelle demeure en soi un problème complexe [Langran,1993], [Egenhofer et Golledge, 1998], [Peuquet, 2002]. Elle ne se limite pas au choix du modèle de référence (linéaire, boucle, branche). Concernant les faits temporels, [Langran, 1993, p.32] note par exemple: *"The view of temporality shared by most researchers envisions a sequence of states punctuated by events that transform one state into the next. States have duration, and are therefore represented by time intervals. Some conceive of events as instantaneous and thus represent them as points ; others consider them to be truncated but potentially telescoping intervals (for example, [Allen; 1983], [Allen; 1984]. Within the simple framework of discrete but linear time, a number of issues arise."* La modélisation de la dimension temporelle *dans les SIG* est un problème plus complexe encore (puisque les contraintes techniques s'ajoutent à la difficulté conceptuelle), que les modèles proposés en standard par les logiciels de SIG simplifient parfois à l'extrême. Une piste de recherche consiste à effectuer des développements informatiques pour enrichir les modèles existants [Badard, 2000]. Dans le cadre d'applications de type observation, notre approche lui est complémentaire : elle exploite les modèles existants et propose une méthode pour enrichir ces modèles avec des informations qu'ils soient capables de gérer.

A.2.2.3. Etudes des évolutions

Comme pour la dimension spatiale, après avoir évoqué les systèmes de référence temporels, nous proposons de distinguer différentes études d'évolutions. Les recherches en psychologie cognitive qui ont montré

l'importance du modèle linéaire à une dimension (fig. A.2.19a), ont également mis en avant l'utilisation de termes « avant/après » pour décrire les relations temporelles (ex.: « Les difficultés sont derrières nous maintenant », « Cela s'est produit avant ton arrivée »)[18]. Pour [Peuquet, 2002, p.27] ces expressions *"... exemplifies what appears to be two movement metaphors within the directionality of time: one in which time (and events contained within it) is moving past the reference base, and another in which the reference base is moving forward (futurward) past stationary events" (fig. A.2.19b et c)*

Nous suggérons de distinguer les études d'évolutions en fonction de la direction temporelle observée : passé/futur. La différence technique entre les deux types d'analyses est ténue, puisqu'elles mobilisent les mêmes outils. Toutefois leur objet, leur utilisation, leur intégration dans les processus de décision, les distinguent plus fondamentalement :

– les unes sont centrées autour de la collecte de données empiriques, tournées vers la description des phénomènes observées, vers l'élaboration de connaissance basées sur l'expérience,

– les autres sont orientées vers la prévision, l'élaboration de connaissances basées sur l'extrapolation et la simulation.

Les deux sont complémentaires : les savoirs tirés du passé permettent d'établir des hypothèses sur ce que peut être l'avenir, et réciproquement la réalisation des évènements « futurs » servent à valider les modèles construits à partir du passé.

18. [Peuquet, 2002, p.26] rappelle l'expérience de [Clark, 1973], qui montre que la description du temps est basée sur une métaphore spatiale aux propriétés spécifiques : « *In general terms, this temporal space of everyday speech is linear, ordered, and directional. The time axis, then like, x,y, or z, axes, can be given an arbitrary zero point marking some reference base and is asymmetrical about this zero point, as in the sense of past, present and future».*

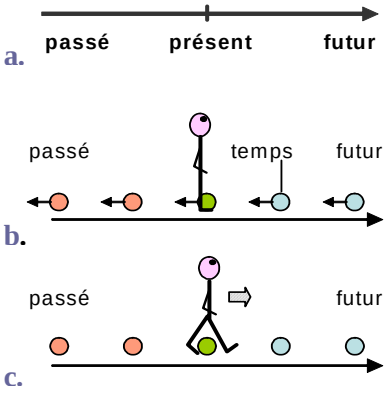


Figure A.2.19. *Perception du temps*
Le temps est généralement perçu de 2 façons : ou bien il défile autour d'un observateur statique (b) ou bien cet individu se déplace dans un paysage temporel (c) [Mangin, 2006]

Analyses descriptives des évolutions passées

Lorsque le regard porte vers le passé, les évolutions ont déjà eu lieu ; il s'agit de les identifier et de les décrire, de les restituer voire de les expliquer, éventuellement de reproduire le phénomène observé en effectuant des études statiques, cinématiques et/ou dynamiques.

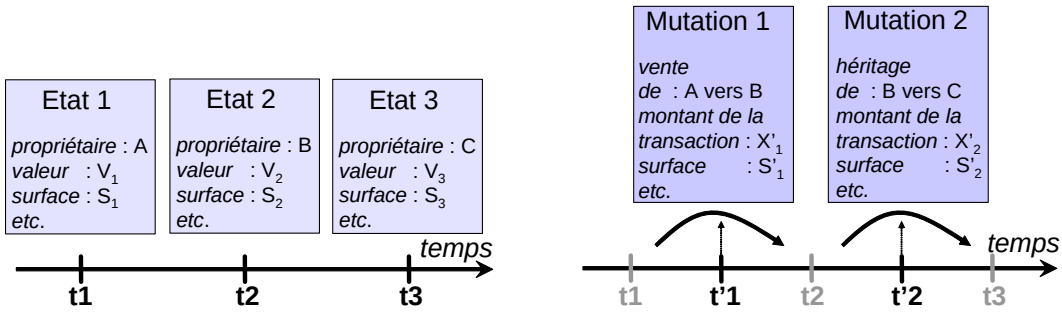


Figure A .2.20. *Exemple de deux descriptions statiques discrètes des évolutions d'un bien immobilier*

La figure de gauche décrit les caractéristiques de faits spatiaux, celle de droite les caractéristiques de faits temporels. Les « mutations » correspondent ici à des événements : les transactions chez le notaire. Elles renvoient aussi aux changements qui résultent de ces événements et qui affectent les faits spatiaux (ex. : changement de propriétaire).

Analyses prédictives, prospectives des évolutions futures

Lorsque le regard porte vers le futur, les évolutions n'ont pas encore eu lieu ; il s'agit de les anticiper. L'étude des évolutions consiste alors à prévoir les états à venir, à projeter des transformations probables pour prédire (voire prévenir) l(es) évolution(s) de l'objet ou du phénomène étudié. Pour cela, il faut un modèle qui permette d'extrapoler sur le futur la situation décrite sur le passé (on utilisera des fonctions d'extrapolation pour

les représentations statiques ou cinématiques, ou un modèle de simulation pour une représentation plus dynamique) : *« il faut bien réaliser ici que aucune prédiction n'est possible sans modèle. Le modèle est au cœur de la prédiction »* [www.larevuedesressources.org]. En supposant que ce modèle reste valide dans le temps (le contexte temporel changeant, le modèle est-il encore valable?) et que l'on soit capable de bien valoriser ses paramètres, on peut élaborer une représentation de la forme future du phénomène.

Dans un cadre posé *déterministe*[19], la pertinence des études prospectives repose entièrement sur la validité du modèle (identification des paramètres, bonne pondération, identification des valeurs seuils, etc.). De sa qualité dépend la qualité des prévisions. Dans un contexte moins déterministe, même avec un modèle de qualité, la validité des simulations se limite souvent à un futur proche. En effet, la démarche de prévision est *« adaptée aux systèmes stables, constitués d'un nombre limité d'éléments, en interactions linéaires (c.a.d pouvant être décrites par des lois mathématique continues et additives). Cependant, les scientifiques se placent de moins en moins dans ce contexte. En effet, bien que la complexité du monde existe depuis toujours, sa reconnaissance est relativement récente. La vision déterministe de scientifiques comme Newton ou Laplace, ne convient plus à l'étude de phénomène complexes organisés, comme ceux rencontrés dans les grands systèmes biologiques, économiques, ou encore* sociaux » [Donnadieu et al., 2003]. Face à ce constat certains concluent à l'impossibilité de prédire l'évolution des systèmes faisant intervenir un très grand nombre d'éléments, à l'impossibilité de connaître de tels systèmes. Ces phénomènes sont alors considérés comme soumis à des interventions du hasard.

19. où l'on suppose que *« tout le futur est (…) entièrement contenu, déterminé par le présent : connaissant les lois du mouvement et les conditions initiales, nous déterminons avec certitude le mouvement futur pour un avenir aussi lointain que nous le souhaitons »* cité par TRI (http://écrits-vains.com/doxa/tri01.htm)

La méthode que nous proposons se place dans le cadre des études descriptives du passé. Ce positionnement nous évite d'avoir à répondre à la question de savoir si l'on peut (et ce que l'on peut) prévoir. Cependant par rapport aux analyses prédictives, nous envisageons ces travaux de deux façons:

– comme étape préliminaire pour la mise en œuvre de prédictions
– comme processus de veille sur les prédictions

La première prend pour hypothèse que la prédiction est possible ; la seconde teste cette hypothèse dans le temps, en envisageant qu'elle puisse s'invalider à tout instant.

– L'observation du passé participe de l'étape préliminaire à toute prédiction ; elle sert à découvrir des régularités, à les caractériser et à les mesurer. Elle précède nécessairement l'étape de modélisation et de théorisation qui consiste à déduire des lois générales des régularités constatées. Ensuite, la prédiction prend pour hypothèse que la régularité du passé se prolonge dans le futur et applique alors au futur les modèles et les lois générales obtenues sur le passé. Dans ces recherches nous proposons de travailler à partir d'informations géographiques modélisées en données, pour effectuer des analyses et des mesures, dans la perspective de trouver des lois qui fonctionnent bien fonctionner sur les données. Il s'agit donc déjà de disposer de données.

– L'observation du passé peut aussi servir dans une démarche de veille sur les prédictions. Les situations observées sont alors des situations du futur proche, devenues présent, puis passé. L'observation établit des indicateurs de type tableaux de bord, pour confronter les résultats des prédictions immédiates à la réalité. Il s'agit moins ici de saisir des régularités que de déceler des irrégularités. Les observations servent alors à :

- compléter la méthode de prédiction, dans un mouvement de feed-back qui introduit de nouvelles informations ; on suppose que le contexte temporel et spatial ne change pas et que les décalages observés ne sont dus qu'aux imperfections du modèle. Ce cheminement analytique en boucle établit le lien entre les analyses descriptives et prédictives.

- suivre l'évolution du phénomène : on suppose ici que les modèles peuvent être fiables, mais que la démarche de prédiction elle-même est incertaine. On observe alors les décalages entre les prévisions et la situation réelle non plus dans l'idée d'améliorer les modèles mais plutôt dans celle de saisir des *bifurcations* ou des *phénomènes émergents*.

A.2.2.4. Les phénomènes dits géographiques

La spécification du caractère évolutif d'un phénomène exprime l'intérêt porté à sa dimension temporelle. Par analogie, dans le phénomène géographique la spécification du caractère géographique renvoie à un intérêt particulier pour la dimension spatiale. Plus précisément, on s'intéresse à la dimension spatiale relativement à un *espace géographique* et pas seulement à un *espace géométrique* [Pumain et Saint-Julien, 1997, p.44].

A.2.2.4.1. Spécificité de la dimension spatiale

La dimension spatiale d'un phénomène peut se réduire à des propriétés métriques et/ou topologiques, mesurées à l'aide d'outils issus de la géométrie. Elle porte sur la distance, la surface, le volume, etc. au même titre par exemple qu'une densité, une concentration ou une masse, en physique ou en chimie. L'espace de ces phénomènes est souvent considéré comme isotrope et sa représentation est généralement celle d'un espace mathématique. Dans certaines recherches, la localisation absolue (en tant que référent à une position sur la terre) ne fait même pas partie des paramètres d'études du phénomène ; seules les localisations relatives (c'est à dire celles des objets les uns par rapport aux autres) sont prises en compte.

L'espace des phénomènes géographiques ne se réduit pas à un espace géométrique euclidien orthonormé. En effet,

– si [Bantels, 1962] (cité par [Peuquet, 2002]) désigne la géographie comme *« the science of distance and space »* exprimant de cette façon l'importance de l'espace comme élément fondamental de la géographie. Pour [Langran, 1993, p.10], citant [Parkes et Thrift, 1980] *«geography differs from geometry, because in geography, space is indivisibly coupled with time »*. Ainsi, *l'espace géographique est évolutif*, ce qui introduit des éléments de complexité dans l'observation des phénomènes géographiques (cf. A.2.2.2). *« The nature of both space and time remains an open issue as the topic of many articles and books in variety of fields »* comme le rappelle [Peuquet, 2002, p.11].

– l'espace géographique est également rarement isotrope. En effet :

- toutes les directions ne sont pas équivalentes (ex. : les jours de vacances, le trafic autoroutier n'est pas le même dans le sens Paris/Province et Province/Paris),
- les régions de l'espace (au sens mathématique) ne sont pas équivalentes (ex. : toutes les destinations ne se valent pas pour passer des vacances « au soleil » en hiver).
- certains phénomènes prennent des formes distinctes et se propagent différemment suivant le lieu où ils se produisent (ex. : le développement urbain ne prend pas la même forme ni la même intensité dans les différentes agglomérations) (cf. figure A.2.21).

Figure A.2.21. *La forme des phénomènes varie en fonction de la localisation géographique. Par exemple, les formes d'urbanisation varient suivant les régions étudiées. Ainsi, on ne trouve pas de favela (à droite) en France.*

Pour les phénomènes géographiques, l'espace ne sert pas seulement de support, il devient une composante à part entière du phénomène. Cette remarque renvoie à la distinction que l'on peut faire entre :

- l'espace comme réceptacle d'objet ou comme place (« topos » pour Aristote),
- et l'espace en tant que : *« système de relation entre des objets matériels signifiant que la localisation dans l'espace est une propriété des objets matériels »* qui stipule que la localisation se distingue de l'objet qui s'y positionne [Petit Robert, 2001].

Ainsi, l'espace géographique ne s'exprime pas seulement en terme de distance ou de forme. Les « portions d'espace » sont désignés par des termes de la langue courantes (ex. : *lieu*, *région* ou *aire)*. Les géographes ont tenté de donner à ces termes des sens précis et limités. Ils désignent des « portions d'espace » qui possèdent un ensemble de caractéristiques propres qui les spécifie (ex. : un *lieu* touristique, une *région* rurale, une *aire* de chalandise). Par rapport aux phénomènes en général, l'étude de la dimension spatiale d'un phénomène géographique conduit donc :

- non seulement à étudier sa répartition ou comment il se «positionne» dans l'espace géométrique,
- mais aussi à analyser en quoi les caractéristiques de l'espace-territoire influent sur cette répartition.

Ce que l'on pourrait résumer rapidement par : l'aspect spatial des phénomènes géographiques est fonction du temps et de la localisation. Cependant, ce n'est pas tant la localisation elle-même, qui distingue deux lieux. C'est la localisation en ce qu'elle fait référence à travers le lieu à l'ensemble de ses caractéristiques (topographiques, sociologiques, écologiques, etc.), caractéristiques qui, elles, influent sur le phénomène. La localisation (x,y,(z)) établit un lien entre l'espace géodésique des phénomènes géographiques et ses différentes propriétés thématiques (propriétés qu'il est courant d'appeler *sémantiques* dans les SIG [http://seig.ensg.ign.fr], cf. A.1.2.4.2).

A.2.2.4.2. Dimension « sémantique »

Les dimensions spatiale et temporelle d'un phénomène géographique sont en relation avec une troisième classiquement qualifiée de *sémantique* dans le domaine des SIG. Ainsi, nous considèrerons avec [Berry, 1964], [Bullock et al., 1974)], [Haggett et al., 1977] ou [Dargemond, 1984], qu'un espace géographique possède trois dimensions (fig. A.2.22). Nous avons déjà vu que l'étude conjointe de la dimension spatiale et de la dimension temporelle est complexe, ce qui conduit généralement à privilégier l'une des deux. L'ajout d'une troisième dimension accentue encore cette nécessité d'une prise en compte différenciée de chacun des aspects. [Sinton, 1978] va plus loin : ayant noté l'incapacité des représentations à

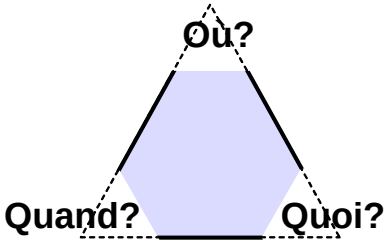


Figure A.2.22. *les trois dimensions d'un phénomène géographique*

intégrer les trois composantes de l'information géographique, il observe qu'en fonction des applications et de la problématique à résoudre, celles-ci possèdent des statuts différents :

- *« one is fixed at a constant value,*
- *a second is controlled to a range of value,*
- *and only the third can be measured on an interval or ratio scale ».*

[Langran, 1993, p.12] met en œuvre ce constat dans une grille pour l'analyse des représentations en données géographiques (fig. A.2.23).

	Fixed	Controlled	Measured
Soils data	time	attribute	location
Topographic map	time	attribute	location
Census data	time	location	attribute
Raster imagery	time	location	attribute
Weather station reports	location	time	attribute
Floods tables	location	time	attribute
Tide tables	location	attribute	time
Airline schedules	location	attribute	time
Moving objects	attribute	location	time

Figure A.2.23. *The representation of geographic data in various formats (extended from Sinton 1978 and from Langran 1993)*

Exemple : Dans une carte topographique, la dimension temporelle est fixe (la date retenue correspond à la mise à jour), les attributs (la composante sémantique) sont contraints par les spécifications, en revanche l'emprise au sol et la localisation des entités topographiques sont mesurées (sur le terrain, par des procédés photogrammétriques) pour définir la composante géométrique.

Exemple: Dans une image raster, la dimension temporelle est fixée (elle correspond à la date de prise de vue de l'image). La localisation est contrainte par le maillage de l'image, en revanche, la valeur attribuée à chaque pixel varie en fonction du niveau de couleur mesuré.

Dernier exemple : Dans les rapports météorologiques, la localisation est fixée (elle correspond à la station météo), la dimension temporelle est contrainte (elle dépend des relevés), en revanche la valeur de température ou de la pression peuvent varier dans un intervalle de valeurs

Remarque : Dans la figure, les données d'occupations du sols (Soils data) sont caractérisées par :
- *un temps fixe (correspondant à la date de mise à jour),*
- *une localisation mesurée (les parcelles d'occupations des sols ne sont pas déterminée)*
- *et des attributs contrôlés (ils prennent généralement leur valeur parmi un nombre fini de codes spécifiés par la légende).*

Pourtant la base MOS ne suit pas ce modèle. En effet, pour permettre d'effectuer des études d'évolutions, elle ne fixe pas le temps. Elle se contente d'en faire une variable contrôlée (les dates de références correspondent aux différentes de mises à jour). Les attributs sont contrôlés (ils prennent leur valeurs parmi les 83 postes de légende). Et la localisation est également contrôlée. En effet pour pouvoir suivre une parcelle dans le temps, la même partition est utilisée pour décrire chaque version d'occupation des sols. Toutefois, cette partition n'a pas été fixée au départ, car elle n'aurait pas pu convenir pour décrire les différents états, elle est adaptée à chaque mise à jour et répercutée ensuite sur les états antérieurs.

Le suivi d'évolutions implique que le temps ne soit pas fixé. La solution adoptée par le MOS pour pouvoir observer des évolutions de l'espace géographique (géométrique et sémantique) a été de renoncer à mesurer l'une pour contrôler les deux.

Au-delà de la formalisation elle-même, cette typologie aide à identifier la dimension principale (« measured ») de chaque étude de phénomène. Réciproquement, elle aide à définir des modèles adaptés aux analyses en spécifiant le rôle de chaque dimension. *«Sinton's framework is important because representational constraints limit the topics and methods of geographic research … A greater challenge is to fix neither location nor time ; to describe the path of a moving object would achieve that goal by fixing attribute (i.e ; the object's identity), but the true challenge lies in not fixing any of the three components of geographic data »* [Langran, 1993]. Cette formalisation nous sert à réfléchir aux études d'évolutions de phénomènes géographiques pour répondre aux questions telles que : « En quoi les études d'évolutions sont-elles des études de phénomènes spatio-temporelles-sémantiques particulières ? », « Qu'est ce qui spécifie les études d'évolutions (de phénomènes géographiques) par rapport aux études (de phénomènes géographiques) en général ? », etc.

A.2.2.3.3. Spécificités des études d'évolutions de phénomènes géographiques

Une évolution est : *« une série de transformations successives »* [Larousse, 2001]. Le caractère successif implique un ordre (avant, après) qui dans le cas des phénomènes géographiques se réfère au temps. Ainsi, l'étude des évolutions d'un phénomène géographique, consiste à étudier des transformations en fonction du temps. Selon la formalisation de [Sinton 1978], le temps serait la variable dite contrôlée. Restent alors à définir les variables fixées et mesurées. Or, aucune des deux composantes, sémantique ou spatiale, ne semble pouvoir être considérée comme fixe. Par exemple, un bâtiment peut se transformer « sémantiquement » (ex. : changement d'activité de bureaux en logement, de commerce d'alimentation en commerce vestimentaire) et/ou «spatialement» (ex. : agrandissement, division). Ainsi, les études d'évolutions de phénomènes géographiques constituent un défi : celui de ne fixer aucune des composantes.

Les propos de Langran cités précédemment *« to describe the path of a moving object would achieve that goal by fixing attribute (i.e ; the object's identity)… »*, nous donnent indirectement des éléments de solutions, en introduisant la notion d'*identité de l'objet*. Si l'on remet en cause l'équivalence faite ici entre *«fixing attribute»* ET (i.e) *« the object's identity »*, nous pouvons alors considérer qu'un objet peut garder son identité malgré des changements de valeurs d'attributs (ou encore que des changements peuvent s'opérer en permettant à l'objet de rester « suffisamment lui-même » cf. chapitre B.3). L'identité de l'objet devient une autre composante de l'information géographique et l'on peut alors compléter le schéma de la figure A.2.22. avec la notion *d'objet* [Peuquet, 2002, p.27] (fig.A.2.24). Alors,

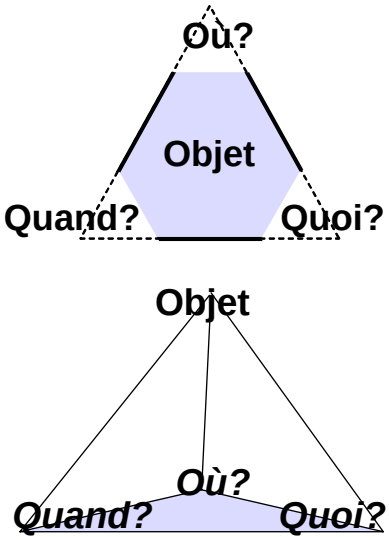


Figure A.2.24. *L'objet et les trois dimensions d'observation d'un phénomène géographique*

- en posant l'identité de l'objet (l'objet est la *variable fixée*),
- on peut choisir de suivre des changements spatiaux et sémantiques (attributs et localisation deviennent des *variables mesurées*),
- en fonction du temps (*variable contrôlée*).

Et l'étude des évolutions d'un phénomène géographique devient alors l'étude de la dimension sémantique et de la dimension spatiale en fonction de la variable temps, l'objet étant fixé.

A.2.3. Vers une méthode d'observation des phénomènes géographiques à l'aide de SIG

Dans le chapitre suivant nous posons les limites des modèles temporels dans les SIG, en particulier nous constatons la difficulté de suivre des mutations et l'impossibilité de saisir des événements avec les modèles utilisés en standard par les logiciels – SIG. Partant de ce constant et en nous inspirant de méthode d'observation en géographie, nous proposerons dans la deuxième partie de construire une méthode d'observation dont les objectifs sont :

- *enrichir les observations statiques* des *évolutions passées* du phénomène (chapitre B.2),
- et surtout permettre à ces observations statiques de *devenir cinématiques*, en simplifiant le *calcul automatique des mutations* (chapitre B.3).

cette méthode devant ainsi permettre d'effectuer des études d'évolutions de phénomènes géographiques où :

- le temps est une *variable contrôlée* (qui dépendra des dates de mises à jour des bases de donnés utilisées),
- la dimension spatiale est une *variable mesurée* (dans la mesure où le thématicien choisira l'objet géographique observé),
- la dimension sémantique est une *variable mesurée* (qui prendra la forme des indicateurs produits par la méthode pour caractériser l'objet observé),

Une fois l'objet géographique à observer déterminé par le thématicien, nous considérons son emprise au sol comme une entité géographique (*portion de territoire*) et nous la prenons alors comme *variable fixe*[20]. Nous montrerons que de façon fondamentale, cela nous permettra d'assurer le suivi automatique des mutations. En effet, d'une certaine façon la méthode que nous proposons généralise le principe des observations temporelles thématiques classiques ; celui-ci consiste à s'appuyer sur la stabilité dans le temps d'une partition de référence pour lui attribuer les différentes valeurs prises par une caractéristique (comme la ligne d'un tableau dans lequel on a crée des colonnes pour accueillir les différentes dates de valeur). Un exemple typique est l'observation des évolutions du recensement par commune : la partition stable est le découpage communale auquel on attribut les différentes valeurs des variables recensées. Cependant,

- les données décrivant les partitions classiques ne sont pas forcément adaptées pour observer les phénomènes [Ribardière, 2005]. Par exemple, le nombre de pavillons dans une commune ne permet pas de savoir combien il y en a dans une zone pavillonnaire particulière.
- et les données décrivant un phénomène, lorsqu'elles existent ne se rapportent pas forcément à une partition stable ; il est alors souvent difficile de les suivre dans un SIG.

On peut alors s'inspirer du MOS. En effet, sa modélisation utilise une partition adaptée à l'observation du phénomène dont elle assure rétroactivement la stabilité dans le temps. Mais cette partition et sa stabilité ne conviennent que pour le suivi des quatre-vingt trois valeurs d'attributs présélectionnées (les attributs sont bien une variable contrôlée)[21].

Dans cette thèse, nous proposons une méthode qui permette de suivre des évolutions d'un phénomène aussi simplement que si l'on disposait d'une partition stable dédiée et sans figer pour cela les attributs observés.

20. Ce qui revient à observer des objets géographiques *non modifiables* (cf. p.55).

21. *The National Historical Geographic Information System* développé aux Etats-Unis par l'équipe de {MacMaster, 2006} offre un autre exemple intéressant. L'objet de ce système est de proposer une base de données nationale des recensements qui soit à la fois multi niveau (il prend en compte les différents niveaux de recensement : des counties jusqu'aux blocs) et multi temporelles (il intègre l'ensemble des données issues de tous les recensements depuis 1790). L'objectif est de permettre aux utilisateurs d'effectuer des observations spatiales (de comparer différentes localisations, d'observer différents niveaux), mais également des observations temporelles (de comparer des états différents)/ Comme pour le MOS, cette base de données est conçue dès l'origine pour observer des évolutions.
Si les derniers recensements américains s'appuient sur une partition commune (situation traditionnelle des bases de recensements), ce n'est plus le cas lorsque l'on remonte plus avant dans le temps, les territoires anciens étant trop différents des territoires actuels. Le travaux de {MacMaster, 2006} consistent à profiter du fait que les modifications sont essentiellement des densifications, pour rétroactivement, tel le MOS, propager les sous partitionnements de la partition - support la plus riche (la plus récente) vers les données les plus anciennes. Après une réallocation des données attributaires au sein des sous-partitions, l'utilisateur disposera alors d'un support unique qui lui permettra d'effectuer des suivis d'évolutions au travers des trois derniers siècles. Cependant, dans cet exemple comme pour les autres recensements, les unités géographiques observés sont des unités administratives qui conviennent à des données de recensement, mais pas forcément à n'importe quel phénomène géographique.

A.3. Observation temporelle dans les SIG

A.3.1. Observations temporelles et SIG

A.3.1.1. Observation de la dimension temporelle

Pour observer la dimension temporelle dans un SIG, les utilisateurs opèrent souvent comme avec des cartes papier : Ils affichent la base de données dans deux états différents et compare deux à deux les 'objets' des deux « images » (comme dans le jeu classique des « sept erreurs », à la différence près qu'il ne sait pas d'avance le nombre de changements à trouver) (fig. A.3.1)

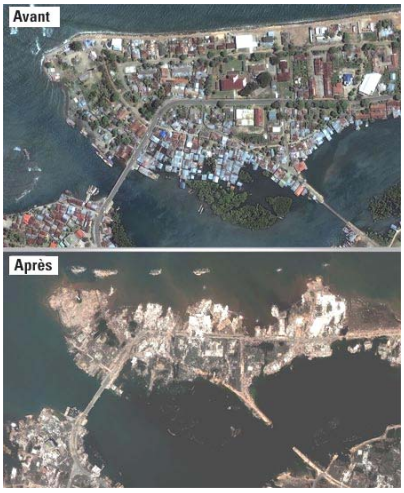


Figure A.3.1. *Images satellites du port de Banda Aceh (Indonésie) avant et après le tsunami de décembre 2004*

Lorsque ces images sont au format *raster*, l'utilisateur peut exploiter un logiciel dit de « traitements d'images », pour tenter une comparaison automatique. Cependant, les conditions de prises de vues, la saison, l'heure, etc. variant généralement d'une fois sur l'autre, les changements de couleurs ou de niveaux de gris sont parfois mal interprétés. Certains travaux de recherches tentent de dépasser ces problèmes [Huet et al, 1996], [Jamet, 1998], [Tellez et Servigne, 1997]. Malgré des résultats prometteurs [Jung, 2002], l'extraction automatique des évolutions n'est pas encore considérée comme totalement opérationnelle de l'avis même de chercheur expert en traitement d'images {Boldo, 2005} (en particulier, il est encore difficile d'extraire automatiquement les évolutions du bâti). En revanche certains autres travaux, qui portent sur des niveaux d'observation géographique plus généraux obtiennent des résultats effectifs, comme par exemple [Weber et Hirsch, 2000] pour les études sur l'évolution de l'urbanisation.

Lorsque les « images » sont au format *vecteur*, ce qui sera le cas tout au long de cette thèse, l'utilisateur bénéficie du format « objet » de ce mode de représentation (cf. paragraphe A.1.2.4.1) : il peut utiliser des requêtes simples pour sélectionner et afficher uniquement les 'objets' (c'est à dire les représentations sous forme de données des *objet*s) qui l'intéressent parmi l'ensemble des 'objets' contenus dans la base, ce qui allège d'autant son observation. Dans la figure A.3.2 par exemple, l'utilisateur concerné uniquement par le 'réseau routier', peut choisir de n'observer que ce type d''objet' au lieu de comparer deux états comportant l'ensemble des données de la BD TOPO®.

Figure A.3.2. *Identification par observation visuelle des changements intervenus entre deux états de la BD TOPO® sur la totalité de la base (en haut) sur un extrait portant uniquement sur le réseau routier (en bas)*

Cependant, même dans ces conditions, le travail d'identification des évolutions reste long et fastidieux, et sans garantie de qualité (d'exhaustivité). Par ailleurs, pour identifier les *changements sémantiques* intervenus sur un type d''objet' particulier, il est nécessaire d'observer

autant de cartes thématiques qu'il y a d'attributs pour ces 'objets'. Ainsi pour le 'réseau routier' de l'exemple précédent, pour identifier les changements sémantiques, l'utilisateur devra réaliser *x* couples de cartes pour pouvoir observer les changements sur les *x* attributs (ex. : le nombre de voies, le classement de la route, le sens de circulation,…) décrivant le 'réseau routier' dans la BD TOPO® (fig. A.3.3).

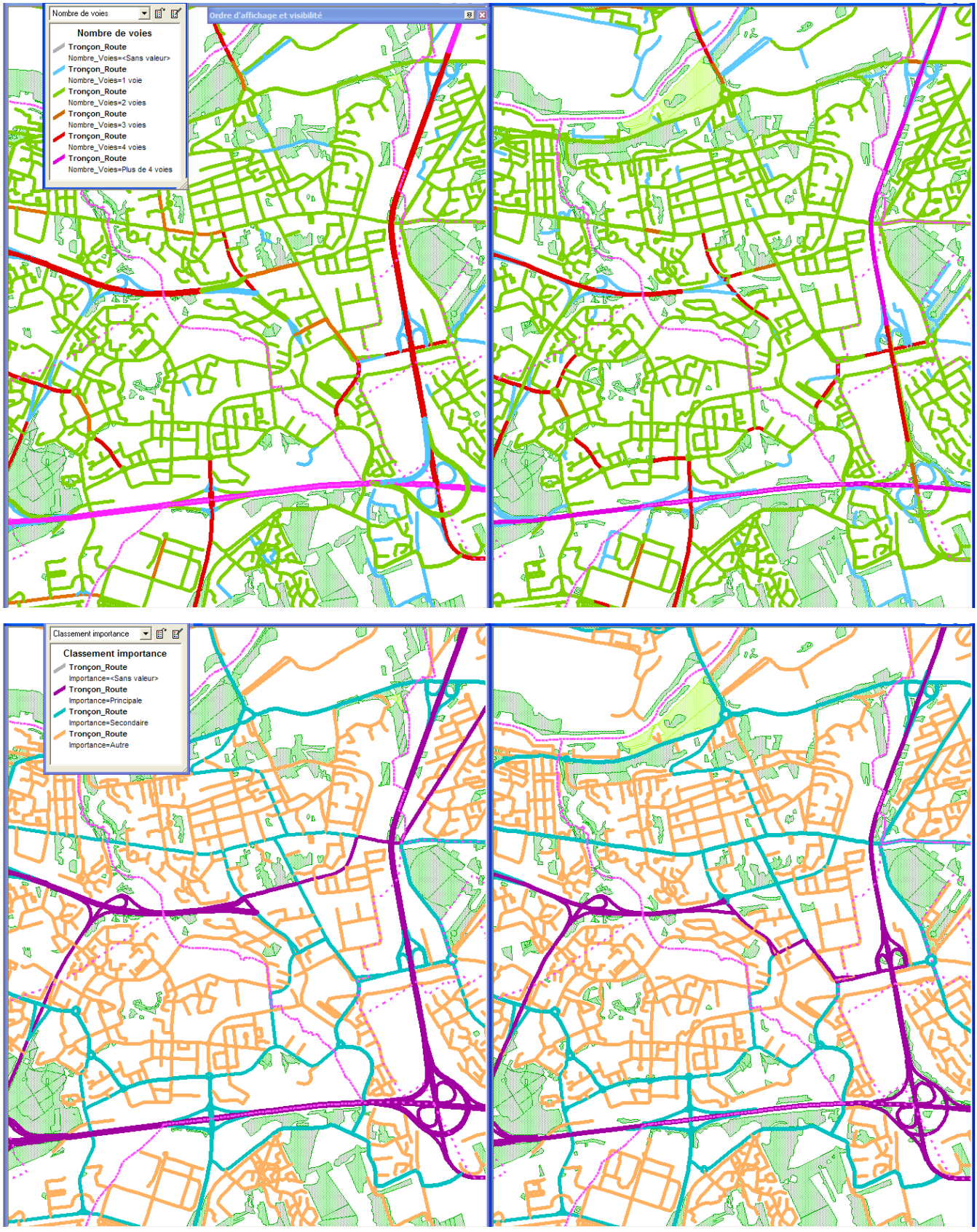


Figure A.3.3. *Identification visuelle des changements sémantiques intervenus sur les 'tronçons de routes' de la BD TOPO®*

A.3.1.2. Observation de phénomènes géographiques

Le suivi d'un phénomène géographique implique de pouvoir prendre en compte la dimension temporelle dans l'observation (cf. paragraphe A.2.1.3.). La qualité d'une observation de phénomènes géographiques *dans un SIG* dépend donc de sa capacité à intégrer cette dimension. Or, les SIG dédiés aux informations *spatiales*, sont mal adaptés à la gestion et à l'analyse des informations *spatio-temporelles*. Ainsi, [Langran,1993, p.4] note : *« Some aspects (applications) are impossible stricly because they can be perform only by temporal GIS, which does not currently exist»*, tandis que [Peuquet,2002] constate :*« Part of this issue (the spatio-temporality) involves the representation of space-time dynamics – something that historically has been ignored in GIS but has always been integrated in process models devoted to and design for certain specific tasks»*. Toutefois, les travaux de recherches fournissent des éléments de solutions. La spatio-temporalité fait ainsi régulièrement partie des thèmes principaux au sommaire des récentes conférences sur les SIG[1]. Malgré ces résultats, l'observation dans un SIG d'un phénomène géographique en mouvement reste une application délicate pour la plupart des utilisateurs. Cela explique probablement pourquoi beaucoup se contentent d'une observation visuelle.

A.3.2. Intégration de la dimension temporelle dans un SIG

A.3.2.1. Le temps dans les SIG : point de vue conceptuel

Le problème du temps dans les SIG résulte de la nécessité de lier simultanément la dimension spatiale et la dimension temporelle. Cette difficulté n'est pas propre aux SIG[2], mais la prépondérance accordée à la dimension spatiale entraîne une modélisation plus sommaire de la dimension temporelle. *« Modern geography has emphasized space rather than time, being rather than becoming, the objective rather than the subjective. This is especially true of spatial analysis and its high-tech offspring , GIS »*. [Couclelis, 1998, p.109]. Dans la pratique, peu de logiciels SIG prévoient autre chose qu'un attribut de type 'date' pour intégrer le temps. Ainsi, la plupart des utilisateurs s'appuie implicitement sur la représentation classique du temps linéaire (au sens littéral) orienté, discret pour situer une suite d'*états* (un état correspondant à la configuration à moment donné de l'ensemble des caractéristique décrivant l'objet) [Couclelis, 1999], [Langran, 1993]. Pour beaucoup d'utilisateurs, ce modèle possède une qualité (en dehors de sa simplicité) : il étend aux bases de données le modèle opérationnel traditionnellement utilisé en cartographie classique, et permet ainsi des analogies (les *versions* des bases correspondent aux *éditions* des cartes[3]). En contrepartie, cette simplification conceptuelle tend à prolonger la perception originelle des SIG comme outils de cartographie numérique. Or, cette perception a pour effet de cantonner les SIG à être plus des cartothèques un peu sophistiquées que de véritables outils d'analyse. *« The limitation of current data models in GIS are due in a large part to the continuing use of the (traditional) cartographic modeling »* {MacMaster, 2006}.

Pour intégrer le temps dans un SIG, il faut passer par une modélisation, qui simplifie alors le problème des relations spatio-temporelles pour les rendre accessibles aux requêtes spatiales, à l'analyse et à la visualisation [Ott et

1. *CoMoGIS2006* (International Workshop on Conceptual Modeling for Geographic Information Systems) 6-9/11/2006 Arizona.
SeBGIS'06 (International Workshop on Semantic-based Geographical Information Systems), 29/10-3/11/2006, France.
DMGIS'05 (Workshop on Dynamic and Multi-dimensional GIS), University of Glamorgan, 5-8/9/2005,Wales, UK.

2. Dans une formulation proche de celle de [Couclelis, 1998] qui note que *« Though treated separatly, Aristote time and space are intimaly connected through motion »*, nous dirions que *« Though time and space are intimately connected through motion, they are treated separatly »*.

3. Le premier modèle de mise à jour (dit par *snapshots*) s'inspire ainsi directement de la mise à jour des cartes (ou des photographies aériennes) : une nouvelle version (ou cliché) pour chaque mise à jour ; le temps ne sert qu'à estampiller les différentes versions indépendantes de la même base de données.
Le lien originel des SIG avec la carte est aujourd'hui encore très fort. Il continue à influencer beaucoup d'utilisateurs dans leur façon d'appréhender les SIG, de concevoir leurs applications et ce, même si la cartographie n'est pas leur objectif principal. Cette analogie qui soutient l'utilisateur débutant dans le domaine de la géomatique, l'accompagne parfois lorsqu'il est plus expérimenté. Ainsi [Langran, 1993] dans son ouvrage de référence sur le temps dans les SIG, « Time in Geographic Information Systems », s'appuie également sur la métaphore cartographique et fait référence la « cartographic time ».

4. Il est a noter dès à présent que la plupart des modèles utilisés dans les SIG, n'intègre pas les *évènements* en tant que tels. Les événements sont les faits temporels. Ils correspondent aux changements entre états. Ne possédant pas de modélisation explicite, ils ne peuvent être saisis qu'à travers leurs effets. Leurs causes, leurs origines ne peuvent pas être intégrées.

Swiaczny, 2001][4]. La mise en œuvre technique et informatique dans les SIG, s'appuie sur

– des modèles de gestion des mises à jour (A.3.2.3.1),

– et sur des méthodes et outils d'extraction des changements (A.3.2.3.2).

Cependant cette simplification introduit des difficultés supplémentaires pour celui qui veut observer un phénomène géographique avec un SIG. De façon un peu caricaturale, on peut en effet affirmer que : le temps des informaticiens n'est pas celui des thématiciens. L'implémentation du temps dans les SIG ne correspond pas exactement à l'observation de phénomènes géographiques, et plus généralement, l'observation de données sur un phénomène n'équivaut pas à l'observation du phénomène lui-même. *« In order to create a truly time integrative GIS, the focus has to change from spatial to temporal and from analysing* changes between events *to the analysis of* change itself» [Ott et Swiaczny, 2001]. En effet, l'opération de mise à jour des données consiste à introduire des changements dans la base de données pour refléter au mieux la réalité du terrain. Elle ne reflète pas *exactement* les évolutions de ce terrain. Elle peut par exemple introduire des changements de type correction d'erreurs qui ne correspondent à aucune modification dans la « réalité ». Par ailleurs, suivant le niveau de sophistication du modèle utilisé pour gérer les changements de la base de données (dit modèle de gestion des mises à jour) l'utilisateur observera des changements plus moins radicaux (certains modèles ne saisissent de modifications que sous la forme de création / destruction, d'autres envisagent des modifications plus subtiles). Poser l'observation des phénomènes géographiques *avec un SIG* revient à postuler que le décalage entre

- l'observation géographique souhaitée
- et l'observation réalisée avec les données,

(qui résulte du passage des *objets géographiques* vers les données géographiques, dans des processus de modélisation successifs) sera compensé par une aide à la gestion ou par la création d'informations nouvelles sur le phénomène.

A.3.2.2. La mise à jour : point de vue technique

Les thématiciens et les techniciens ont des points de vues différents sur la dimension temporelle dans les SIG (fig. A.3.4):

- pour les thématiciens, il s'agit de pouvoir suivre des *objets géographiques*, d'analyser des évolutions et d'étudier le temps comme composante du phénomène observé.

- les techniciens, quant à eux, considèrent le temps plutôt comme un paramètre externe qui affecte la qualité des données qu'ils administrent[5]. Il leur impose d'introduire régulièrement des changements dans ces données, au risque sinon pour les bases de donnée de devenir obsolètes. L'intégration de la dimension temporelle est alors principalement une opération de maintenance (temporelle) sur les bases de données. Cette opération est appelée la *mise à jour*.

D'un point de vue technique, la mise à jour permet d'établir un lien entre l'observation des données et celle du phénomène géographique étudié. Cependant, elle n'équivaut pas à l'introduction de la dimension temporelle ; elle demeure fondamentalement technique. Des définitions mettent en avant cet aspect. Ainsi [Dell'Erba et Libourel, 1997], [Spéry, 1999, p.77-80] la présente sous forme d'évolutions de *schémas*[6] et/ou de données, rappelant que les changements introduits dans les données ne résultent pas uniquement d'évolutions sur le terrain. Ils peuvent également correspondre à des aspects techniques de maintenance et de gestion des données. Il existe plusieurs types de mises à jour, de

Technicien	Thématicien
Donnée	Objet géographique
Changements	Évolutions
Le temps, comme source d'obsolescence	Le temps, comme paramètre d'étude, comme composante

Figure A.3.4. *Différences de points de vue entre techniciens et thématiciens*

5. L'actualité fait partie des critères servant à qualifier la qualité d'une base de données, tout comme la précision géométrique ou sémantique, l'exhaustivité [David et Fasquel, 1997]. Des données qui ne décrivent plus la réalité du monde étudié, ne permettent plus aux SIG qui les utilisent d'aider à la décision. Les données en perdant ainsi de leur intérêt, perdent alors également de leur valeur –marchande– en particulier.

6. Un *schéma conceptuel* de données, est une structure graphique permettant de présenter le contenu d'une base de données: les 'objets', leurs *attributs*, les *relations* entre 'objets'

maintenances techniques. Certaines n'agissent que sur les données, d'autres interviennent également au niveau de la structure de la base de données elle-même. Il s'agit soit :

– de mises à jour intervenant sous la forme de correction d'erreurs, qui introduisent des changements au niveau des données. Elles sont difficiles à identifier, (à moins de (pouvoir) comparer à une autre source de données et de constater que l'autre base comporte les mêmes modifications, ex. : comparaison avec une photographie aérienne), car elles apparaissent comme des évolutions « vraies ». Ce sont soit

- des **corrections d'erreurs vraies**, par exemple elles compensent l'oubli d'un 'objet géographique', elles rectifient la valeur d'un attribut ou la forme d'une emprise, etc. Après ces mises à jour la qualité des données est meilleure.

- ou des **corrections dues à des différences d'interprétations** des *spécifications de saisie.* Ces mises à jour n'améliorent en général pas la qualité des données. En effet, dans certains cas ambigus (liés par exemple à l'usage de termes comme « objet remarquable », « centre d'intérêt », « important », etc.) ou du fait d'un défaut d'interprétation, l'opérateur de saisie des données peut décider d'introduire des corrections. Pensant introduire une « vraie » mise à jour ou une correction d'erreurs, il introduit alors un changement d'une autre nature. Par exemple, nous avons cru déceler un phénomène d'urbanisation spécifiquement centrée autour des garages en zone pavillonnaire, alors qu'il ne s'agissait que d'une différence d'interprétation des spécifications (la figure A.3.5 montre comment nous est « apparu » le « phénomène »).

– de mises à jour en introduisant des modifications dans la structure des données (en plus de celles sur les données elles-mêmes). Ces mises à jour sont généralement plus faciles à repérer (la disparition d'un attribut, l'apparition d'un nouvel 'objet' ou d'une nouvelle couverture de zone se voient clairement). Ces mises à jour regroupent :

- les **changements de spécifications**. La décision du producteur des données de modifier les *spécifications* de la base de données[7], implique des changements dans la base de données « pour la mettre à niveau », c'est à dire pour la rendre conforme aux nouvelles règles qu'elle doit respecter.

- les **enrichissements de la base**. Ces enrichissements peuvent prendre la forme d'une introduction de données supplémentaires, par exemple l'ajout d'un simple attribut sur un 'objet géographique' (ex. : le trafic sur le 'réseau routier') ou d'un nouvel 'objet' (ex. : les 'sentiers piétonniers'), voire d'un nouveau *thème* (ex. : l'hydrographie, la végétation). Ils peuvent également prendre la forme d'une extension de la zone couverte par la base.

A.3.2.3. La gestion des mises à jour : point de vue pratique

La mise en œuvre de mises à jour soulève de nombreux problèmes. Beaucoup portent sur des aspects organisationnels concernant la collecte, l'intégration ou encore la livraison des mises à jour[8] [Raynal, 1997], [Bordin, 2002, p.168-172]. Nous ne présenterons ici que les aspects de modélisation et d'implémentation. Pour mieux illustrer les enjeux de ces modèles pour l'observation de phénomènes géographiques, dans ce paragraphe nous appréhenderons la mise à jour de façon restrictive. Nous adopterons une définition –proche de celle des utilisateurs thématiciens–, qui ne considère que des modifications en rapport avec des évolutions du terrain physique :

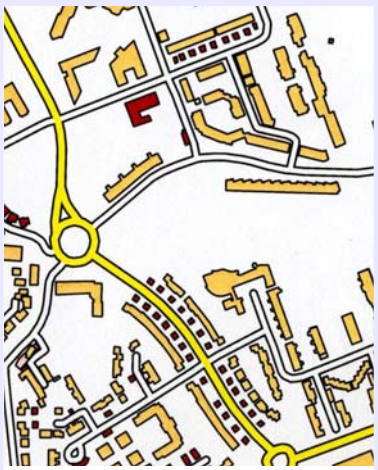


Figure A.3.5. *Des données nouvelles ne correspondent pas forcément à des informations nouvelles*

Expérience issue du cas d'étude sur les *zones pavillonnaires* (cf. Partie C) :

Alors que nous étudions l'évolution des *zones pavillonnaires*, nous avons eu l'occasion d'observer *l'apparition de structures nouvelles constituées de petits bâtiments alignés (représentés en rouge dans l'extrait ci-dessus), nous laissant croire un temps qu'un phénomène urbain nouveau se produisait. Une photo-interprétation temporelle à partir de clichés plus anciens a montré que ces bâtiments étaient des garages et qu'ils avaient toujours existé ! Ces 'bâtiments' présentés comme de nouveaux 'bâtiments' dans la base de données correspondaient en fait à du bâti déjà anciens…*

7. qu'il s'agisse des spécifications :
- *de contenu* (ex. : regroupement de deux attributs en un seul qui les résument)
- ou *de saisie* (ex. : modification de la surface minimale requise pour qu'un objet géographique surfacique soient représentées sous forme de polygone).

Pour aider les utilisateurs a identifier ces modifications, il importe de constituer des *méta-données* (données sur les données) bien renseignées, en particulier pour les changements de spécifications de saisie, qui ne sont pas explicitées dans la base de données.

8. Pour chacun de ces aspects, plusieurs formes de solutions se présentent généralement au technicien-producteur. Par exemple, pour la collecte il doit décider la façon dont il souhaite l'organiser (sous forme d'une unité dédiée ou en unités décentralisées ou encore s'il essaie d'établir des partenariats). Pour l'intégration, il lui faut définir le rythme de la mise à jour (tous les 5 ans, tous les ans, en continu, etc.), les unités d'intégration (ex. : un département, une région ou tout autre unité), etc. ; Pour la livraison, il doit déterminer les formats de livraison, le type d'informations livrées (ex. : livraison différentielle, journal des modifications).

La *mise à jour* est alors l'opération qui tend à ajuster les données à l'actualité de la situation sur le terrain. Elle prend en compte les évolutions des *objets géographiques* et les traduit sous forme de changements sur les données et fait évoluer ainsi la base de données pour assurer sa pérennité.

A.3.2.3.1. Modèles d'implémentation et de gestion des mises à jour

Il existe plusieurs modèles pour intégrer les changements effectués sur les données d'une base de données [Badard, 2000, p.23]. Ils possèdent des qualités différentes en terme :

- de commodité d'accès aux différents états temporels d'un même 'objet' au sein d'une base de données,
- d'explicitation des liens existant entre ces différents états temporels, pour aider à établir la série temporelle de ces états,
- de caractérisations de ces liens décrivant les changements opérés, il s'agit ici de posséder des informations sur le changement lui-même, comme par exemple son type (création/destruction/modifications) ou sa date d'occurrence, voire son origine, sa cause).

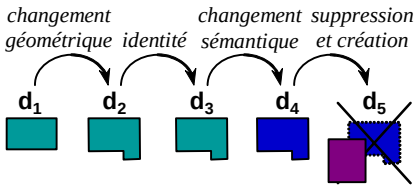


Figure A.3.6. *Exemple de série temporelle de données géographiques*

Cet exemple montre les différents états d'un 'objet' créé à la date d1, suite à une série de changements (respectivement un changement géométrique à la date d2, pas de changement à d 3, un changement sémantique à d4, une destruction avec création d'un nouvel objet à d5).

Dans ce qui suit nous avons choisis de présenter quatre modèles. Il ne s'agit pas d'un état de l'art. Nous souhaitons seulement montrer la difficulté d'observer des informations temporelles dans un SIG et illustrer combien ces observations dépendent des modèles de gestion mis en œuvre dans la solution SIG. Les quatre modèles seront présentés en fonction de la richesse d'informations qu'ils fournissent sur les changements de données et de la facilité à reconstituer la suite des états successifs d'un « même » 'objet'. Cette présentation suit un ordre croissant relativement bien corrélé avec la complexité de mise en œuvre. Pour illustrer les différences, nous utilisons le cas particulier de la figure A.3.6. Les modèles que nous avons retenus sont :

Le *modèle par snapshots* qui gère la mise à jour des bases de données comme une mise à jour de carte papier : à chaque mise à jour correspond une version de la base. Chaque version est indépendante des autres, il n'y a aucun lien entre les données des différentes versions (fig. A.3.7a).

Le *modèle par time stamp*, qui archive différents états d''objets'. Dans ce modèle la base de données peut être vue comme une « donnéethèque » sorte d'équivalent d'une cartothèque pour les données ; c'est une collection de représentations d''objet géographique'. Deux représentations différentes sont indépendantes ; même si elles correspondent dans la « réalité » à deux états d'un même *objet* ; il n'existe pas de lien formalisé entre elles (fig A.3.7b.)

Le *modèle par versionnement* d'objet qui archive différentes *versions* d''objets'. Ce modèle considère qu'un même 'objet' peut avoir plusieurs états et formalise un lien (par le biais d'un numéro de version) entre les différentes versions de cet 'objet'. En revanche pour éviter les informations redondantes, chaque version ne comporte qu'une description partielle de l'état de la donnée correspondant aux caractéristiques modifiées. (fig A.3.7c).

Le *modèle par historique* qui introduit le concept de *successeur*. Chaque 'objet géographique' est décrit dans sa totalité. S'il change d'état, il existe alors un lien formalisé par le biais d'un identifiant pour spécifier l'état qui lui « succède » (fig A.3.7d.).

Dans le détail de ces différents modèles, nous montrerons en particulier comment les améliorer pour obtenir des informations sur la série temporelle des états d'un « même » 'objet'. Nous verrons aussi comment

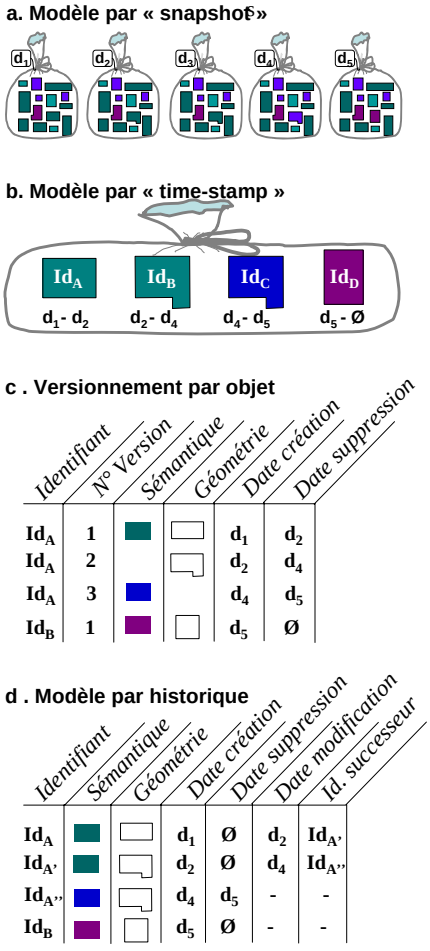


Figure A.3.7. *Différents modèles pour intégrer les mises à jour*

faciliter les requêtes temporelles élémentaires telles que *« Quels sont les 'objets' présents à la date d ? Quels sont ceux supprimés après d ?, etc.»*.

Le modèle par *snapshots* ou par estampillage de la base de données

Le modèle par *snapshots*, qui reprend la métaphore de la prise de vue instantanée, est le modèle le plus élémentaire (cf. figure A.3.8). Dans ce modèle, la mise à jour d'une base de données est conçue comme un processus ponctuel[9], pendant lequel l'ensemble des données est « mis à jour » avec pour résultat une *base de données à jour*. Cette base, sorte de représentation statique (ou d'instantané) à une date d fixée (dite *date de mise à jour*), est archivée en tant que *version* de la base à la date d. La livraison des « mises à jour » consiste alors à fournir des duplicata de cette base archivée. Les différentes versions de la base se succèdent alors, sans qu'il existe de lien automatique entre les différentes versions, celles-ci étant conçues comme des *jeux de données* autonomes[10].

La simplicité du modèle explique son succès. Conceptuellement, il correspond au modèle classique d'observation discrète, statique (que nous expérimentons régulièrement, ne serait-ce qu'en regardant un film). Pour les utilisateurs habitués aux cartes papiers, il transpose le mode traditionnel de mise à jour en cartographie. Enfin pour les utilisateurs de SIG, il adapte pour la dimension temporelle une structuration proche de celle déjà adoptée par certains logiciels de SIG pour gérer la dimension spatiale (ex. : les *couches* de Mapinfo, les *thèmes* d'ArcGis, les *calques* d'Autocad) (cf. Annexe 2). Pratiquement, il est simple à implémenter, puisque la *gestion des évolutions* se limite à leur collecte et à leur intégration sous forme d'interventions pour modifier les données ; il n'y a pas de *gestion des données d'évolutions* (c.a.d. on ne traite pas de questions telles que : « Q*uelles sont les données qui ont changé ? En quoi ont-elles changé ? Pourquoi ont-elles changé?*»). Le modèle par *snapshots* n'implique pas d'étudier préalablement comment modéliser, collecter, exploiter, analyser la dimension temporelle : Il est le plus simple à livrer et le plus simple à recevoir.

En revanche, sa simplicité limite ses performances pour le suivi de phénomènes dans un SIG. Il ne comporte en effet aucune *donnée d'évolution*, c'est à dire aucune information sur les modifications effectuées sur les données, qui pourraient servir de « trace » d'évolutions réelles. En fait, ce modèle offre le minimum d'information temporelle : il ne propose pas d'autre information que la *date d'estampillage* associée au jeu de données[11]. Il arrive que cette date ne soit même pas implémentée dans les données. Elle n'apparaît alors que sous la forme d'une inscription sur le support ou le bon de livraison ! Dans ce cas, si un champ (ou attribut) de type « date d'estampillage » n'est pas préalablement introduit et automatiquement rempli avec la bonne valeur et ce, pour les différentes versions de la base de données étudiées, alors lorsque l'utilisateur affiche simultanément deux versions pour pouvoir comparer visuellement deux états différents de la base, il n'a aucun moyen de distinguer de quelle version sont issues les données. Le modèle par *snapshots* peut toutefois être amélioré simplement et à moindre coût en introduisant un *identifiant* ou une *signature*[12]. Cette opération qui consiste à créer un attribut supplémentaire à chaque 'objet', permet alors par simple requête de sélectionner les 'objets' dont l'identifiant se trouve dans une version antérieure mais est absent des versions suivantes (destruction) (et réciproquement les 'objets' dont l'identifiant est absent d'une version antérieure mais présent dans celles qui suivent (création)) .

9. Le processus « ponctuel » de mise à jour s'oppose conceptuellement à un processus de mise à jour considéré comme « continu », où les informations peuvent être intégrées à tout moment dans la base de données et non pas seulement aux dates correspondant à la version de la base. Dans la pratique, la production effective de ces mises à jour *ponctuelles* dure généralement plusieurs mois.

10. Un *jeu de données* est un extrait d'une base de données, conçu comme un tout, la mise à jour et la livraison de données pouvant ne concerner qu'une partie d'une base de données plus globale.

11. Techniquement, cette *date d'estampillage* peut correspondre à des dates différentes : date de livraison, date de la version de la base, date du début (ou de fin) de la mise à jour (dans le cas où la mise à jour n'est pas en continu),etc. qui ne correspondent pas exactement à la date « d'apparition » sur le terrain.

12. La définition de l'identifiant en informatique peut être :
- *stricte,* on parle alors de signature : On change d'identifiant dès que l'objet subi un changement. Ce qui revient à considérer que tout objet ayant subi une modification n'est plus identique à lui-même, qu'il a disparu pour être remplacé par un autre. La gestion de cet identifiant (ou signature) peut alors être sous traitée à l'ordinateur de façon automatique ;
- *souple*, on parle alors d'identificateur : l'identifiant est lié à l''objet'. Tant que le producteur de données estime que l''objet' reste semblable à lui-même malgré les modifications, il lui associe la même valeur 'd'identificateur'.

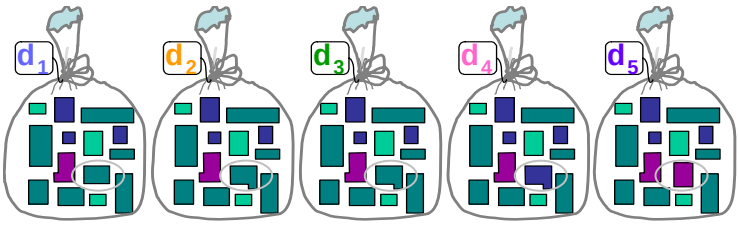


Figure A.3.8. *Principe du modèle par* snapshots

Le snapshots *archive les différentes versions de la base de données de façon autonome. Sans identifiant, les jeux de données temporels sont indépendants. Rien ne permet d'identifier automatiquement un 'objet' dans la base de données. Il faut le repérer et le suivre visuellement dans les différentes versions.*

Le modèle par *time stamp* ou par estampillage des données

Si on enrichit le modèle basique par *snapshots* en ajoutant pour chaque 'objet' un identifiant strict (signature) et deux champs supplémentaires : 'date de création' et 'date de suppression'[13] [Langran et Chrisman, 1988], [Langran, 1993], on peut alors intégrer dans une même base de données l'ensemble des données géographiques issues des différentes mises à jour. Cette base de données unique regroupe alors toutes les données (passées et actuelles) décrivant un même territoire : nous sommes alors dans le modèle par *time stamp*. La date de création et de suppression étant attachées à chaque 'objet', on peut alors envisager des dates qui ne soient plus liées à des versions de carte mais qui soient des dates ayant une signification plus proches des *objets* eux-mêmes. Ce modèle n'implique en effet pas une mise à jour ponctuelle et sait au contraire profiter d'une mise à jour en continu. Ce modèle permet facilement de connaître les objets présents ou absents à une date fixée, ainsi que la durée de vie d'un *objet* particulier (ex. :*« Quels sont les objets dont la valeur de l'attribut 'date de création' est inférieure à d et celle de l'attribut 'date de destruction' supérieure à d*'?») (cf. figure A.3.9a). Si on intègre par ailleurs un identifiant « souple » ou identificateur (c'est à dire un identifiant susceptible d'être utilisé par deux données distinctes), si on considère qu'elles sont la représentation de deux états distincts d'un même objet (cf. note 12), alors on aura un lien entre deux versions différentes du même objet. Ce modèle permet éventuellement de saisir des évolutions plus élaborées que les simples création/suppression (cf. figure A.3.9b). Techniquement, il implique une gestion d'un volume important de données, mais reste simple à implémenter.

13. La définition des dates peut varier. Par exemple pour une autoroute, la date de construction peut être : la date où la construction est actée, la date où le fuseau (ou le tracé) définitif est décidé, la date où la chaussée est terminée, la date d'ouverture au public ou encore la date d'introduction dans la base de données.

Le versionnement par objet

Dans le modèle précédent, chaque état d'un même 'objet' est enregistré comme une donnée avec sa période de validité donnée par le couple {'date de création' ; 'date de suppression'}. Si l'on choisit de changer d'identifiant dès que cet état subi un changement (identifiant strict), alors les différents états d'un « même » 'objet géographique' sont indépendants les uns des autres. Il n'existe pas de relation entre eux. Si l'on choisit de conserver le même identifiant souple (identificateur) pour les modifications qui préservent l' « identité » de l'objet, alors on peut avoir différents états d'un même objet. On peut reconstituer les liens de succession par requêtes sur la date de création (ou la date de suppression) (ex. : *«Sélectionner les objets dont l'identifiant est XXX et trier les par ordre croissant en fonction de la 'date de création'* »). Dans ce cas, le lien de succession n'est pas directe : pour l'obtenir il faut effectuer des requêtes spécifiques. Le modèle par *versionnement* propose quant à lui de formaliser ce lien, en introduisant un

numéro de version. Pour obtenir la série temporelle des états, il suffit alors de faire les chaînages des différents numéros de versions, à valeur d'identifiant constante. Ce modèle a été conçu pour réduire le nombre de données archivées et pour éviter d'avoir à gérer un volume important d'information (comme c'était le cas dans le modèle précédent) ; il optimise le stockage des données en ne dupliquant pas les données redondantes. Ainsi, la description de chaque version d''objet' ne comporte que les informations nouvelles (modifiées), mais pas celles qui n'ont pas changées (cf. figure A.3.9c). La reconstitution de l'état complet d'un 'objet géographique' met en œuvre un mécanisme de gestion interne à ce système qui permet de remonter les informations au travers des différentes versions. Dans la pratique, pour obtenir la série temporelle des différents états d'un 'objet', il est nécessaire de reconstituer la description de chaque état[14]. Ce qui dans la pratique est fastidieux. L'exploitation de ce modèle peut être simplifiée en acceptant d'introduire des informations redondantes et d'associer à chaque version la description entière de l'état de l''objet'. On obtient alors un modèle très proche du modèle par *time stamp* : seul le numéro de version les distingue.

14. Pour cela il faut réaliser une succession d'opérations : d'abord il faut trouver la dernière version (« sélectionner l''objet' dont l'identifiant est XXX et tel que le numéro de version est au maximum V_m»), puis établir les changements avec la version précédente V_{m-1} et intégrer les éléments de l'état complet que V_{m-1} contient, s'intéresser à la version V_{m-2}, établir les changement entre V_{m-2} et V_{m-1}, introduire les éléments d'état général contenu dans V_{m-2} dans V_{m-1} et V_m, et s'intéresser à V_{m-3}, …et ainsi de suite jusqu'à la première version V_1. En reprenant en sens inverse, la suite des états ainsi reconstitués, on obtient alors les états successifs de l''objet' observé.
Cette reconstitution demeure toutefois assez lourde. Ainsi, bien que ce modèle permette de reconstituer automatiquement des évolutions d''objets géographiques', les utilisateurs lui préfèrent souvent le modèle par *snapshots* plus sommaire ; Ils trouvent plus simple d'effectuer des analyses visuelles, car ce modèle est conçu pour limiter le volume de données stockées, mais au détriment de la facilité d'accès aux différents états des données.

Les quatre pages suivantes constituent les quatre parties de la figure A.3.9.

Figure A.3.9. *Comparaisons de modèles de gestion des mises à jour et saisie des changements.*

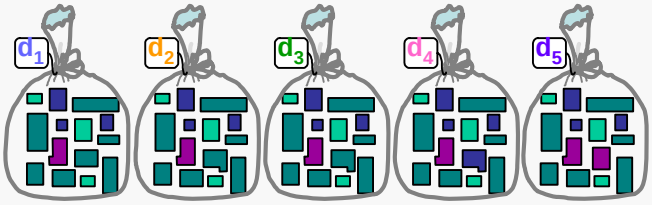
Cette figure tente de résumer dans une présentation unifiée, les avantages et inconvénients des quatre modèles présentés. La présentation du modèle par historique lui fait suite.

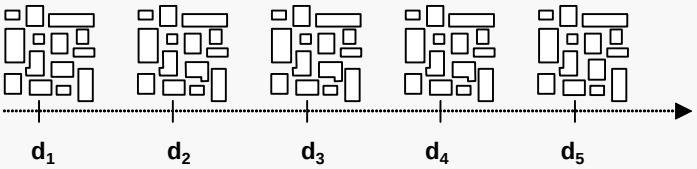
Modèles et variations

Données et états des données

a1. « Snapshot » sans attribution de la date d'estampillage générale aux données

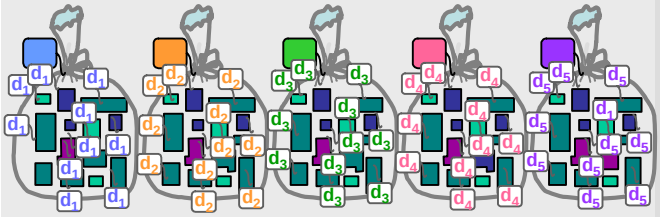
⇒ *Plusieurs versions de la même base*

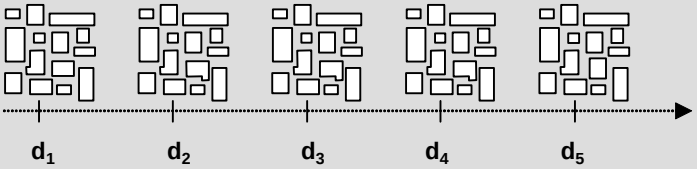




⇒ Une date est affectée à chaque ensemble de données.
Rien ne permet d'identifier automatiquement un objet dans la base de données. Il faut le repérer et le suivre visuellement.

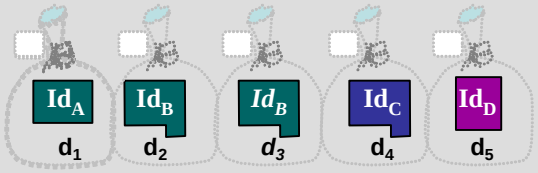
a2. « Snapshot » avec attribution de la date d'estampillage générale aux données

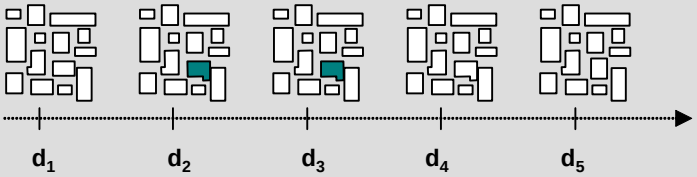




⇒ La date globale de la version est affectée à chacune des données. Rien ne permet d'identifier automatiquement un objet dans la base de données. Il faut le repérer et le suivre visuellement.

a3. « Snapshot » avec signature

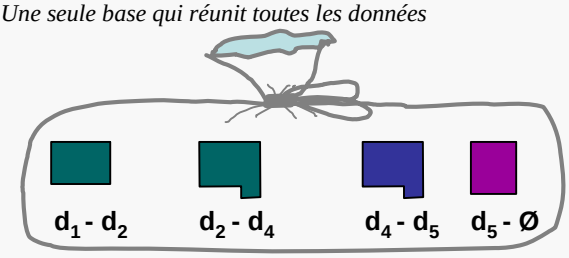


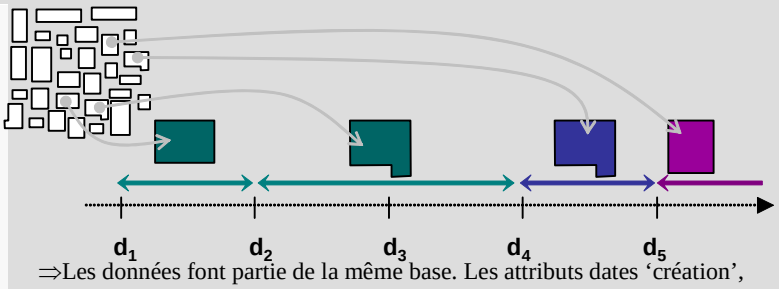


⇒ La signature permet de distinguer un état particulier et donc d'identifier ceux qui sont identiques d'une version à l'autre.

b1. « Time stamp »

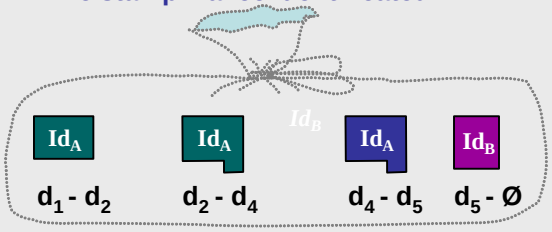
⇒ *Une seule base qui réunit toutes les données*

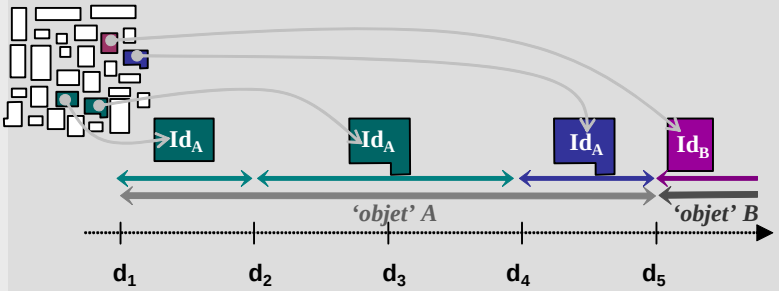




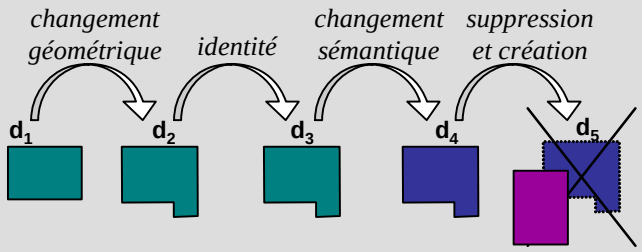
⇒Les données font partie de la même base. Les attributs dates 'création', 'suppression' permettent de connaître la période de validité d'un état d''objet'.

b2. « Time stamp » avec identificateur





⇒L'identificateur permet d'identifier automatiquement les différents états d'un même objet dans le temps. Les attributs dates de 'création', 'suppression' permettent de connaître leur période de validité. Et de reconstituer la suite des états.



Identification des changements
Reconstitution de la série des états

a1. « Snapshot » sans attribution de la date d'estampillage générale aux données

⇒ *Plusieurs versions de la même base*

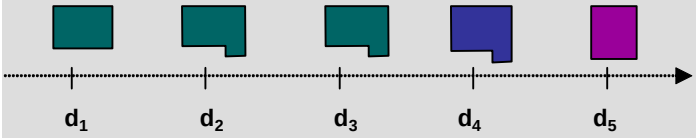
** La série des états d'un 'objet' ne peut pas être reconstituée automatiquement. Les données sont indépendantes les unes des autres*

** On ne peut pas saisir les changement.*

** Si on intègre toutes les données dans une même base, on ne distingue pas de quelles versions ces données sont issues.*

** Ce modèle ne peut servir que pour une observation visuelle «manuelle» ou pour des applications où la dimension temporelle n'importe pas.*

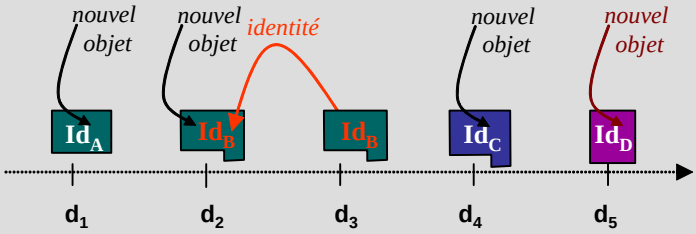
a2. « Snapshot » avec attribution de la date d'estampillage générale aux données



** La série des états d'un 'objet' ne peut pas être reconstituée automatiquement. Les données sont indépendantes les unes des autres*

** Pour une date correspondant à une version on a facilement l'état de la base à cette date en sélectionnant les 'objets' par l'attribut : 'date d'estampillage'*

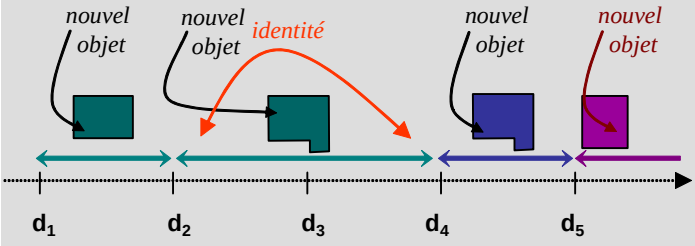
a3. « Snapshot » avec signature



** La série des états d'un 'objet' ne peut pas être reconstituée automatiquement (car dès que l''objet' est modifié, il change de signature). On ne distingue pas les données créées/supprimées des données modifiées.*

** On identifie simplement les 'objets' qui n'ont pas changé entre deux versions: ce sont ceux dont la signature se retrouve parmi les signatures des deux versions .*

b1. « Time stamp »

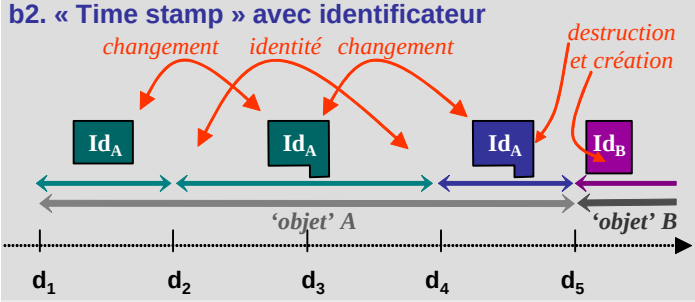


** La série des états d'un 'objet' ne peut pas être reconstituée automatiquement, car chaque état d'objet correspond à une donnée indépendante.*

**Pour une date d fixée, on identifie les nouveaux objets: ce sont ceux dont la' date de création' est d (pour les objets détruits, il faut que 'date de suppression' soit égale à d). En revanche, on ne distingue pas les objets créés/ supprimés des objets modifiés.*

** Pour un état d''objet' fixé, on accède directement à sa période de validité par requête sur le couple: (date de création; date de suppression).*

b2. « Time stamp » avec identificateur



** La série des états d'un 'objet' peut être reconstituée automatiquement, en sélectionnant à identificateur constant, les 'objets' par ordre croissant de 'date de création'.*

** Les 'objets' créés sont ceux pour lesquels il n'existe pas d''objet' ayant à la fois le même identificateur et une 'date de création' antérieure à la leur (les 'objets' supprimés sont ceux pour lesquels il n'existe pas d''objet' ayant à la fois le même identificateur et une 'date de suppression' postérieure à la leur).*

** Pour un état d''objet' fixé, on accède directement à sa période de validité par requête sur le couple: (date de création; date de suppression).*

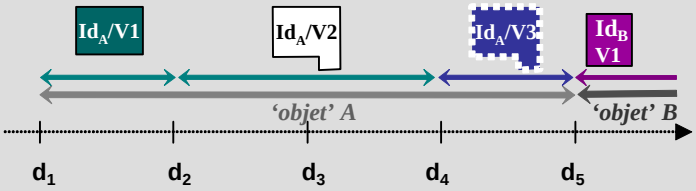
Modèles et variations

Données et états des données

c1. Versonnement d'objet sans description totale

⇒ *Une seule base avec différentes versions d''objet'*

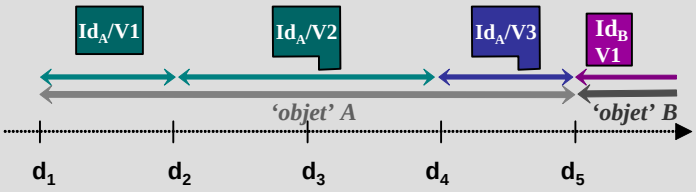
Identifiant	N° Version	Sémantique	Géométrie	Date création	Date suppression
Id_A	1	■	□	d_1	d_2
Id_A	2	-	⌐	d_2	d_4
Id_A	3	■	-	d_4	d_5
Id_B	1	■	□	d_5	Ø



⇒A chaque état d''objet' sont associés: l'identificateur de l'objet représenté et un numéro de version. Pour établir la suite des états, il faut retrouver les versions successives ainsi que la valeur des attributs qui n'ont pas été modifiés. Les attributs dates 'création', 'suppression' permettent de connaître la période de validité de cet état.

c2. Versionnement d'objet avec description totale

Identifiant	N° Version	Sémantique	Géométrie	Date création	Date suppression
Id_A	1	■	□	d_1	d_2
Id_A	2	■	⌐	d_2	d_4
Id_A	3	■	⌐	d_4	d_5
Id_B	1	■	□	d_5	Ø

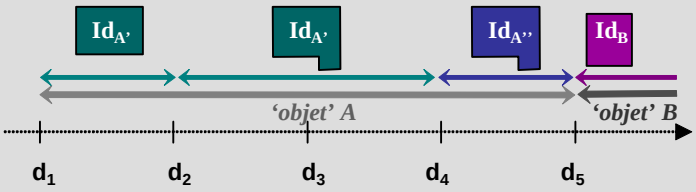


⇒A chaque état d''objet' sont associés: l'identificateur de l'objet représenté et un numéro de version. Pour établir la suite des états, il faut retrouver les versions successives. Les attributs dates 'création', 'suppression' permettent de connaître la période de validité de cet état.

d1. Historique

⇒ *Une seule base qui réunit toutes les données*

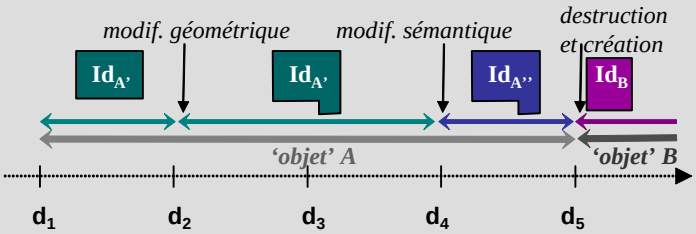
Identifiant	Sémantique	Géométrie	Date création	Date suppression	Date modification	Id. successeur
Id_A	■	□	d_1	Ø	d_2	$Id_{A'}$
$Id_{A'}$	■	⌐	d_2	Ø	d_4	$Id_{A''}$
$Id_{A''}$	■	⌐	d_4	d_5	-	-
Id_B	■	□	d_5	Ø	-	-



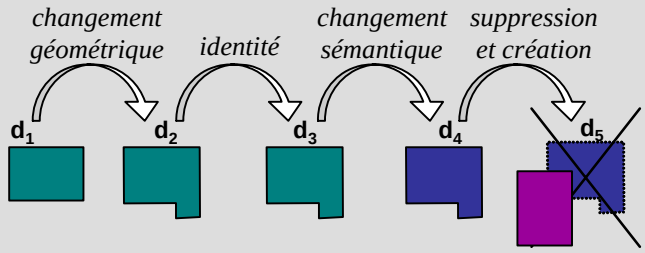
⇒Chaque état d''objet' a sa propre signature. Le lien entre les différents états doit être retrouvé par requête sur l'attribut 'id. du successeur'. Les attributs dates 'création', 'suppression' permettent de connaître la période de validité de cet état.

d2. Historique et informations différentielles

Identifiant	Sémantique	Géométrie	Date création	Date suppression	Date modification	Id. successeur	Type modification
Id_A	■	□	d_1	Ø	d_2	$Id_{A'}$	chgt. géom.
$Id_{A'}$	■	⌐	d_2	Ø	d_4	$Id_{A''}$	chgt. sém.
$Id_{A''}$	■	⌐	d_4	d_5	-	-	destruct°
Id_B	■	□	d_5	Ø	-	-	créat°

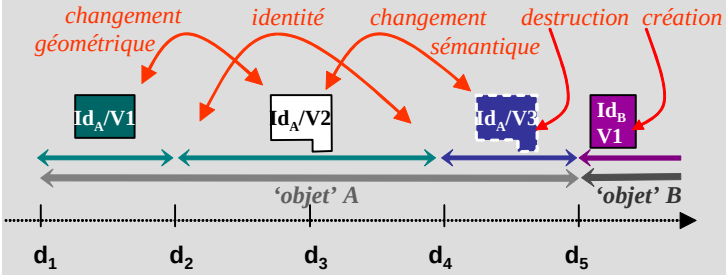


⇒Chaque état d''objet' a sa propre signature. Le lien entre les différents états doit être retrouvé par requête sur l'attribut 'id. du successeur'. L'attribut 'type de modification' explicite le changement subi par l'état de l''objet'. Les attributs dates 'création', 'suppression' permettent de connaître la période de validité de cet état.



Identification des changements
Reconstitution de la série des états

c1. Versonnement d'objet sans description totale



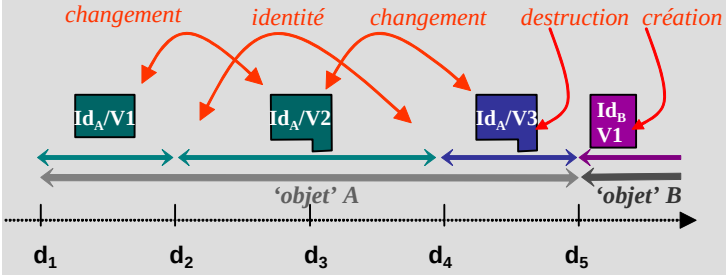
** Une série des états* partiels *d'un 'objet' peut être reconstituée automatiquement, en sélectionnant à identificateur constant, les 'objets' par ordre croissant de 'numéro de version': La description de chaque état ne comporte que les valeurs d'attributs qui ont été modifiés.*

** Les 'objets' créés sont ceux dont le numéro de version est 1. Les 'objets' détruits sont ceux dont pour lesquels la 'date de suppression' est non vide et le 'numéro de version' est maximal*

** Pour un état d''objet' fixé, on accède directement à sa période de validité par requête sur le couple: (date de création; date de suppression).*

** A chaque changement de version l''objet' subi un changement (ou une destruction), pour connaître sa nature, il suffit de regarder la(les) composantes et attributs concernés.*

c2. Versionnement d'objet avec description totale



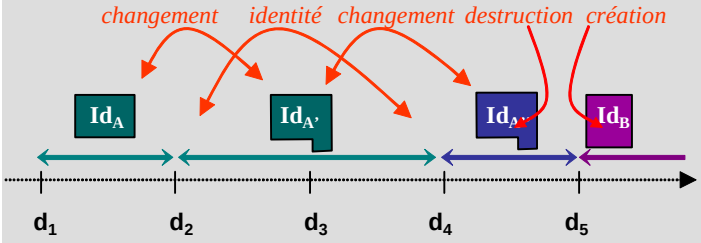
** La série des états d'un 'objet' peut être reconstituée automatiquement, en sélectionnant à identificateur constant, les 'objets' par ordre croissant de 'numéro de version'.*

** Les 'objets' créés sont ceux dont le numéro de version est 1. Les 'objets' détruits sont ceux dont pour lesquels la 'date de suppression' est non vide et le 'numéro de version' est maximal*

** Pour un état d''objet' fixé, on accède directement à sa période de validité par requête sur le couple: (date de création; date de suppression).*

** A chaque changement de version l''objet' subi un changement (ou une destruction), mais il n'y a pas de moyens automatiques pour préciser sa nature.*

d1. Historique



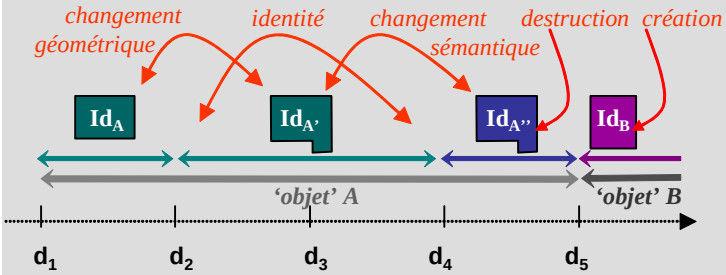
** La série des états d'un 'objet' peut être reconstituée automatiquement, en sélectionnant de proche en proche la série des état successifs à partir de l'attribut 'Id. successeur'.*

** Les 'objets' créés sont ceux dont l'identifiant n'est pas une valeur prise par l'attribut 'id. successeur'. Les 'objets' détruits sont ceux dont la 'date de suppression' est non nulle.*

** Pour un état d''objet' fixé, on accède directement à sa période de validité par requête sur le couple: (date de création; date de suppression).*

** Chaque fois que la 'date de modification' est non nulle il y a eu un changement (sémantique et/ou géométrique), mais il n'y a pas de moyens automatiques pour en préciser la nature.*

d2. Historique et informations différentielles



** La série des états d'un 'objet' peut être reconstituée automatiquement, en sélectionnant de proche en proche la série des état successifs à partir de l'attribut 'Id. successeur'.*

** Les 'objets' créés sont ceux dont l'identifiant n'est pas une valeur prise par l'attribut 'id. successeur'. Les 'objets' détruits sont ceux dont la 'date de suppression' est non nulle.*

** Pour un état d''objet' fixé, on accède directement à sa période de validité par requête sur le couple: (date de création; date de suppression).*

** Chaque fois que la 'date de modification' est non nulle il y a eu un changement (sémantique et/ou géométrique), pour en préciser la nature, il suffit de regarder la valeur de l'attribut 'type modification'.*

L'historique des données

La mise à jour par gestion de l'historique des données est l'une de celles qui s'attachent le plus à l'aspect temporel. Elle ne vise pas seulement la gestion technique des changements dans les données. Elle propose d'établir un lien entre les différents états d'un 'objet' qui soit directement accessible. Elle ébauche ainsi une première solution exploitable par les utilisateurs pour faire du suivi d'évolutions.

Ce modèle est proche du modèle par *time stamp* avec identifiant, mais il introduit en plus la notion de *successeur*. Il associe ainsi systématiquement à tous les objets :
un *identifiant* et les deux attributs 'date de création' et 'date de suppression' mais également une 'date de modification' et un 'identifiant du successeur'.

Lorsqu'une donnée doit être modifiée, il faut alors définir si :

- l'objet géographique représenté par la donnée a disparu et s'il a été remplacé par un autre (ex. : destruction d'un bâtiment et construction d'un nouveau),
- ou si l'objet géographique décrit par la donnée ayant évolué, la donnée n'est plus une bonne représentation de cet objet (ex. : réhabilitation et agrandissement d'une construction).

Dans le premier cas, on remplit le champ 'date de suppression' avec la date à laquelle l''objet géographique' a « disparu » et on crée une nouvelle donnée pour représenter le nouvel 'objet géographique', avec un nouvel identifiant et une 'date de création' correspondante. Dans le second cas, on remplit le champ 'date de modification' avec la date à laquelle l''objet' géographique a « évolué ». On inscrit dans le champ 'identifiant du successeur' la valeur de l'identifiant attribué à la nouvelle donnée créée pour le successeur. Cette donnée reçoit pour valeur de champ 'date de création', la date à laquelle l'objet géographique considéré a évolué (cf. figure A.3.9d).

Dans ce modèle, l''identifiant du successeur' établit un lien direct entre les différents états d'un 'objet géographique' *O*. Il facilite la reconstitution automatique de la suite évolutive de ces états successifs. Pour cela, il suffit d'extraire l''identifiant du successeur' *O+1* de *O*, de le sélectionner par requête grâce à son identifiant, on obtient alors l'état de *O* après sa première évolution, puis de reproduire cette opération autant de fois que *O* évolue, c'est à dire tant que la valeur de la 'date de suppression' reste un champ vide. On notera ici que la valeur de l'attribut 'date de modification' permet d'associer simplement et directement une date à chaque évolution.

Ce modèle qui simplifie le suivi d'évolutions reste techniquement lourd à gérer. Par ailleurs, il est délicat à mettre en œuvre par des producteurs de données qui ne sont pas les utilisateurs finaux de ces données. En effet, l'interprétation des évolutions d'un *objet géographique* est souvent thématique ; Suivant le contexte, les évolutions seront perçues différemment. Ainsi, pour un architecte un bâtiment de bureaux transformé en habitations a évolué en subissant des *modifications*, tandis que pour le service qui recense les entreprises, il a *disparu*. Parfois, même dans un contexte thématique identique, il est difficile de déterminer si l'objet est modifié ou s'il a disparu. (Par exemple dans le phénomène de mitage où dans un mouvement « sournois » des habitations s'installent petit à petit dans un espace vert : *Quand dira -t'on qu'il évolue et à partir de quand dira -t'on qu'il a disparu au profit d'une zone d'habitat ?*). Ces questions sur le maintien de « l'identité » des *objets* au cours de leur évolution (et sous-jacent de leur identifiant dans la base) rend difficile toute

interprétation des évolutions par un autre acteur que l'utilisateur-thématicien. (Dans le chapitre B.3, nous détaillons cette problématique que nous appelons la problématique de *l'objet suffisamment lui-même)*. Il s'agit en effet de préciser dans quelles mesures et dans quelles conditions les évolutions d'un objet n'affectent pas son intégrité, les réponses variant en fonction de l'utilisateur. Dès lors, le producteur qui introduit dans les données une pré-interprétation des évolutions des *objets* (ex. : un 'bâtiment' est '*modifié*' au lieu de '*détruit*') peut certes aider certains utilisateurs à la saisie des évolutions étudiées, mais il risque également de proposer des évolutions qui ne correspondent pas à « la réalité » de certains autres.

A.3.2.3.2. Méthodes et outils d'extraction des changements

Nous avons vu plusieurs modèles pour intégrer les informations de mises à jour. Certains proposent des moyens pour reconstituer automatiquement la série temporelle des états d'un 'objet géographique' ; d'autres se limitent à proposer un ensemble de données indépendantes. Certains considèrent les mises à jour comme les étapes d'un processus d'évolution[15] ; d'autres les intègrent comme des opérations autonomes dont ils ne retiennent que le résultat final, parfois de façon globale (au niveau d'un jeu de données et pas de chaque donnée). Il est important de remarquer qu'aucun de ces modèles ne propose d'informations explicites sur les *changements* intervenus sur les données. Ces modèles qui intègrent la dimension temporelle fournissent uniquement un lien pour relier les *états* avant/après. Ils ne permettent pas de décrire le changement intervenu entre les deux. Des outils et des méthodes issus de développements informatiques spécifiques existent toutefois, qui peuvent venir compléter les modèles d'intégration des mises à jour en fournissant des éléments directs sur les *changements* opérés sur les données. Nous en présentons trois, du plus standard au plus compliqué à mettre en œuvre :

– la clef MD5

– la journalisation

– les méthodes et outils d'appariement et de calculs de différentiels

15. La maintenance par mise à jour est alors vue comme un processus d'évolution portant sur la base de données, visant à refléter au mieux le processus d'évolution se produisant sur le terrain.

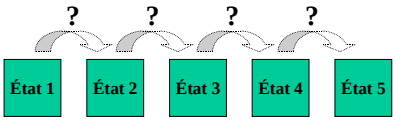


Figure A.3.10. *L'utilisateur peut avoir une description détaillée des états des objets géographiques qu'il observe sous forme de données, mais il ne possède pas d'information sur les événements qui les modifient.*

Il faut noter ici que l'intégration de ces outils et méthodes ne permet toutefois pas de modéliser directement les faits temporels que sont les *événements* : ces méthodes donnent des détails sur les modifications opérées entre deux états de données, mais elles ne gèrent pas d'information directe sur la nature de l'*événement* à l'origine du changement (cf. figure A.3.10). En particulier, même en complétant les modèles de gestion de la dimension temporelle par des outils d'extraction des changements, l'utilisateur ne sera pas en mesure de savoir : *Pourquoi les objets géographiques qu'il observe ont changé, qui (ou quoi) est à l'origine de ce changement*, ou toute autre information détaillée sur cette modification.

La clef MD5

Le premier outil d'extraction des changements présenté, la *clef MD5*, n'est pas spécifique au domaine des SIG : c'est un outil informatique standard. Il permet d'identifier rapidement et simplement les données qui ont changé, même si elles n'ont subi qu'une modification minime (par exemple sur une valeur d'attribut), et ce, sans nécessiter de passer en revue toutes les données et toutes leurs valeurs d'attributs. En revanche, cet outil ne donne aucun détail sur la nature des changements.

La *clef MD5* est un code, une sorte « d'empreinte digitale » pour les messages informatiques. Pour générer ce code, on utilise l'algorithme MD5

qui permet de produire une clef de 128 bits à partir d'un message de taille quelconque. Cet algorithme permet de vérifier l'intégrité des données d'un message de façon beaucoup plus sûre que le classique contrôle de parité. Il sert lors des transferts de messages : l'ordinateur émetteur génère une clef MD5 pour signer le message envoyé, puis envoie le message et la clef au destinataire. A la réception le destinataire re-calcule la clef MD5 du message et la comparer avec celle envoyée par l'émetteur. Si les deux clefs sont identiques, la transmission s'est bien passée, dans le cas contraire le destinataire sait que le message a été altéré durant la transmission et peut éventuellement demander sa réémission.

Dans le cadre du suivi d'évolutions au sein de bases de données géographiques, la clef MD5 peut être utilisée lorsque le modèle de référence est un modèle par *snapshots* basique sans signature. Elle permet d'identifier simplement et automatiquement les données restées identiques entre deux versions. Pour cela, il suffit de calculer la clef MD5 correspondant à chaque donnée géographique pour chacune des versions considérées, puis de comparer les deux listes de clés. On sait alors que les 'objets géographiques' dont la clef MD5 a changé ont été modifiés entre les deux versons de la base. A l'inverse, tout 'objet géographique' dont on retrouve la clef dans les deux listes est resté strictement identique dans les deux versions de la base. On obtient ainsi facilement les *non-changements* ou changements de type *identité*.

La journalisation

La *journalisation* consiste à enregistrer au fur et à mesure l'ensemble des commandes informatiques qui sont effectuées sur une base pour en modifier des données. Partant de l'état initial de cette base, il suffit de lancer la suite des commandes enregistrées dans le journal pour reproduire la suite des opérations qu'elle a subie et obtenir ainsi au final l'état le plus récent de la base. La mise en œuvre d'un tel journal implique certains développements pour saisir, inscrire et gérer les transactions dans un fichier dédié (ou table de stockage). Le système peut être sophistiqué par exemple en sélectionnant les interventions enregistrées (ex. : on n'enregistre pas les opérations de zoom sur la carte qui ne correspondent pas à une modification « vraie » de la base mais plutôt à une opération de confort de visualisation). Le journal donne un accès direct à tous les changements opérés sur la base, mais il reste factuel : il les décrit mais ne donne pas leur cause. Par exemple il peut révéler qu'un 'bâtiment' a été supprimé, mais on ne sait pas si le *bâtiment* participe à un complexe plus large suite à une opération de rénovation ou s'il a été effectivement détruit sur le terrain.

Par ailleurs, le journal ne peut apporter de solution effective que si la base initiale considérée correspond exactement à la base à l'origine du journal. Ainsi, si un utilisateur intervient sur les données géographiques d'un producteur autre, en les modifiant pour intégrer ses propres données, il interrompt la chaîne logique du journal. Il aura alors des difficultés pour intégrer les changements fournis par le journal du producteur initial. Par exemple, si l'utilisateur coupe un 'tronçon de route' en deux 'sous-tronçons' pour prendre en compte un attribut supplémentaire pour lequel le dit 'tronçon' n'est pas homogène et que pendant ce temps le producteur coupe également ce 'tronçon' lors d'une mise à jour pour intégrer un *croisement*, alors il y aura incohérence entre les deux états de la base (celle de l'utilisateur et celle du producteur), ce qui empêchera tout report automatique du changement 'création d'un croisement'.

Les méthodes et outils d'appariement et de calculs de différentiels

Des méthodes pour identifier et qualifier des changements de données.

L'*appariement* consiste à « reconnaître » dans deux jeux de données les couples d''objets' servant à représenter un même *objet géographique* dans chacun des jeux. Appliqué à deux bases de données différentes, cela permet d'établir des relations entre leurs représentations pour faire par exemple de la *multi-représentation* dans un SIG. Utilisé sur deux jeux d'une même base de données, mais dans des versions différentes, l'appariement permet de sélectionner les objets qui n'ont pas la même représentation dans les deux jeux de données et d'identifier ainsi ceux qui ont changé. Pour pouvoir comparer deux données, les outils d'appariement utilisent des *outils de calculs de différentiels*.

Le *calcul du différentiel entre deux données* consiste à expliciter la différence existant entre les deux et à la qualifier. Si ces données représentent deux états différents d'un même 'objet', le calcul du différentiel caractérise alors la différence existant entre l'état-avant et l'état-après. Il sert ainsi à obtenir des éléments explicites sur les changements intervenus sur les données.

Typologie des changements

Les changements identifiés dépendent des conditions posées dans les programmes de calcul, et entre autre de la typologie des changements implémentée dans ces programmes. Certains outils ne proposent d'identifier que des changements élémentaires (création/suppression), d'autres proposent de saisir des changements plus subtils. Ainsi, [Badard, 2000, p.49], définit 8 changements :

- trois élémentaires : création, destruction et stabilité (ou identité)
- deux détaillées : modifications géométrique et sémantique
- trois complexes établissant des liens informatiques n : m entre objets de versions différentes : scission, agrégation, fusion (fig. A.3.11).

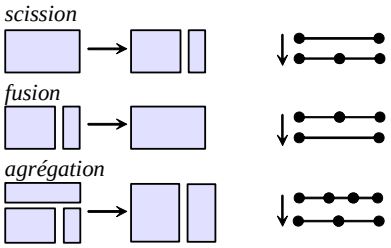


Figure A.3.11. *Exemples de changements complexes (Extrait de [Badard, 2000, p.49])*

Il est à noter ici que les informations sur les changements, étant obtenues de façon *automatique* par des *programmes informatiques* sur des *données*, ne portent de fait que sur des *changements sur des données*. Or les changements sur les données géographiques ne sont pas les *évolutions sur les objets géographiques*. En effet, les changements ne sont que des *représentations d'évolutions* –lorsqu'ils ne correspondent pas, par ailleurs, à des modifications techniques sur la base de données (assurance qualité, changement de spécifications). Cette différence introduit de notre point de vue un décalage fondamental dans le processus d'analyse de phénomènes géographiques avec un SIG.

Mise en œuvre des techniques d'appariement et de calculs de différentiels.

Les méthodes et outils d'appariement et de calculs différentiels sont aujourd'hui développés principalement par les producteurs de bases de données, comme éléments de processus de production : processus de mises à jour [Badard, 2000], processus de mesures de qualité des bases de données [Bard, 2004] ou encore de création de bases de données dérivées [Mustière, 2001]. Ces techniques restent complexes à mettre en œuvre et beaucoup sont de l'ordre de la recherche (ex. : les travaux en appariement de données vecteurs de [Devogele et al., 1998], [Devogele, 1997], [Bel Hadj Ali, 1997], [Bucaille, 1996], [Lemarié et Raynal, 1996]). De fait, elles sont encore peu accessibles aux utilisateurs.

A.3.2.3.3. Les métadonnées

Les modèles pour implémenter les mises à jour ne modélisent pas explicitement les *événements* : il n'y a pas d'information pour les représenter en tant que tels. Les techniques de calculs différentiels et d'appariement peuvent servir à déterminer les changements opérés entre deux états d'une même donnée, mais ils ne saisissent que des informations *directement déductibles* des données. La structure définie par le modèle ne permet en revanche pas d'intégrer d'autres informations caractéristiques (par exemple, la cause du changement : technique (ex. : correction d'erreur, changement de spécifications) ou thématique (ex. : destruction d'un tronçon de route liée à la réalisation d'un rond-point (fig. A.3.12), construction d'un tronçon pour desservir une future zone industrielle), leur durée réelle (la période de temps séparant deux versions de la base peut être d'un an alors que le changement s'est effectué sur un mois, sur deux ans) etc.). L'introduction de *métadonnées* peut pallier ce manque.

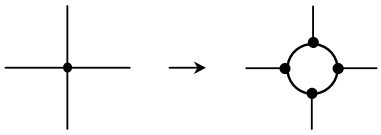


Figure A.3.12. *Transformation d'un carrefour en rond-point : changement ou évolution ?*

Les *métadonnées*, qualifiées souvent de « données sur les données » regroupent, généralement sous forme d'un fichier complémentaire, des données contenant des informations utiles à la description de la base de données ou des données qu'elle contient. Il existe plusieurs niveaux de métadonnées (fig. A.3.13). Au niveau détaillé, on peut introduire des informations sur les changements effectués sur les données pour préciser leurs causes ou leur durée mais aussi qui a collecté l'information ou qui l'a introduit dans la base de donnée, etc.

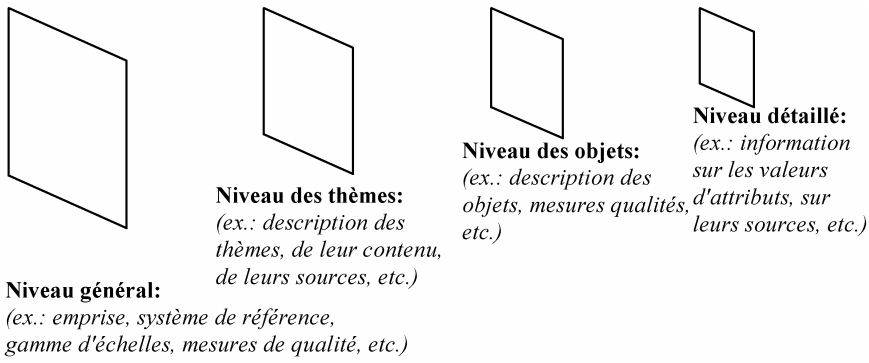

Niveau général:
(ex.: emprise, système de référence, gamme d'échelles, mesures de qualité, etc.)
Niveau des thèmes:
(ex.: description des thèmes, de leur contenu, de leurs sources, etc.)
Niveau des objets:
(ex.: description des objets, mesures qualités, etc.)
Niveau détaillé:
(ex.: information sur les valeurs d'attributs, sur leurs sources, etc.)

Figure A.3.13. *Différents niveaux de métadonnées (extrait de [Bordin, 2002 , p.179])*

A.3.3. Deux approches pour réaliser une observation temporelle dans un SIG

Les modèles d'implémentation des mises à jour, les outils et méthodes pour reconstituer des informations sur les changements et les métadonnées constituent un ensemble de moyens complémentaires pour mieux gérer les évolutions d'une base de données. Introduits dès la conception des bases, ils permettent de saisir des informations sur la dimension temporelle plus nombreuses et plus élaborées. Leur mise en œuvre associée augmente les capacités des outils SIG à intégrer le temps dans les analyses et à le traiter de façon automatique. L'exemple développé dans ce qui suit montre que l'ensemble de ces moyens permet de construire des solutions *techniques* au problème de suivi d'informations spatio-temporelle dans les SIG. Toutefois, il illustre également le degré de technicité requis pour les développer.

A.3.3.1. Approche technique : un exemple de configuration pour gérer la dimension temporelle

Une expérience récente développée à l'IGN[16] offre un exemple de configuration élaborée permettant un suivi opérationnel des changements. Elle atteste des capacités techniques des SIG à devenir des outils d'analyses plus performants. Cette application est bâtie autour d'un *modèle* qui s'inspire à la fois du modèle par *versionnement* par objet *avec description totale* (on y retrouve la notion de *version*) et du modèle par *historique* (il propose 'une date de modification').

16. Il s'agit du projet *Unification des bases de données*, mené par F. Fuchs.

Ce modèle associe à chaque donnée :

- Un identifiant d'*objet*, qui correspond de façon unique à l'objet représenté par la donnée.
- des attributs sémantiques et géométriques, décrivant de façon complète l'*état de l'objet* représenté par la donnée,
- ainsi que trois dates : création, modification et destruction, donnant des éléments sur la validité de l'*état* et de la donnée.

Chaque état d'un *objet géographique* fixé est représenté par une donnée, qui correspond alors à une version différente de l''objet'. Cette version n'est pas explicitée par un numéro de version. Elle est connue indirectement grâce à l'identifiant d'*objet* (qui sert à déterminer quelles sont les données versions du même objet) et à la 'date de création' (qui sert à ordonner ces versions).

Ce modèle introduit également un identifiant de modification, appelé 'numéro de réconciliation', la réconciliation étant l'opération qui crée une nouvelle donnée. Cet identifiant permet de « pointer » vers un fichier de *métadonnées* (dite table des réconciliations) qui contient les informations sur les modifications effectuées sur la base. Chaque réconciliation y possède : un 'numéro' (l'identifiant), et un 'contenu' de la réconciliation décrivant l'ensemble des changements opérés sur les données qui a conduit à la création de la nouvelle donnée. Le champ 'contenu' est rempli automatiquement grâce à un programme développé pour enregistrer « à la volée » les modifications réalisées sur une donnée (outils de *journalisation*), pour les interpréter et les intégrer dans le fichier de métadonnées. Ainsi, l'utilisateur peut accéder directement à :

- une description circonstanciée des changements (stockée dans la table de réconciliations)
- une description complète de chaque version de l'objet (stockée à chaque nouvelle réconciliation dans une table spécifique (dite table d'historique).

Par ailleurs, le 'changement' étant représenté en tant que tel dans cette modélisation (il possède un identifiant et une description propre), on peut lui ajouter d'autre champ (comme par exemple un champ 'origine' de la réconciliation pour signifier si l'origine est technique ou thématique). Ces champs supplémentaires pourraient également servir à spécifier les *évènements géographiques* à l'origine des changements lorsque ceux-ci ne sont pas des « changements techniques ».

Ce modèle est fonctionnel : il facilite l'accès à la série temporelle des états d'un objet ; il discrimine des changements variés (création / suppression mais aussi modifications) ; il fournit directement les informations sur ces changements ; il offre même la possibilité d'introduire des informations supplémentaires sur ces changements pour aider l'utilisateur à mieux les comprendre. Conçu avec un système d'interfaces conviviales adapté à *ses* utilisateurs, pour eux il est simple d'utilisation (en l'occurrence, les utilisateurs principaux du système sont les opérateurs chargés de mettre à jour les données sur le terrain). Cependant, la mise en œuvre d'un tel

système requiert de nombreuses compétences informatiques : en modélisation, en conception, en programmation, etc. Elle implique une expertise spécifique en information géographique pour mobiliser, organiser et développer les ressources des logiciels (SIG et autres systèmes). L'ensemble de ces compétences n'est pas toujours à la disposition des utilisateurs thématiciens. C'est pourquoi, pour proposer une solution pour observer des phénomènes géographiques dans les SIG, nous avons choisi une autre approche, qui ne repose pas sur l'exploitation des capacités informatique des SIG : une approche méthodologique.

A.3.3.2. Approche méthodologique : vers une méthode pour élaborer des données de suivi dans un SIG

Pour proposer une solution au problème de suivi de phénomènes géographiques dans un SIG, nous proposons de revenir au principe de base qu'il faut des données pour utiliser des SIG. Il s'agit donc ici de disposer de données décrivant le phénomène que puissent être observées et suivies dans le temps avec un SIG. Pour répondre à ce besoin, nous proposons de développer une méthode pour en construire, avec pour contraintes sur les données produites de :

- décrire le phénomène en fonction des points de vue thématiques considérés,
- être complémentaires aux données pouvant exister par ailleurs sur le phénomène, pour venir enrichir les observations déjà disponibles,
- pouvoir être observées avec un SIG,
- pouvoir être suivies dans le temps et ce quels que soient le modèle et les outils de gestion temporelle utilisés.

Dans cette thèse, nous ne considérerons pas un phénomène particulier. Nous voulons construire une méthode générique pour observer *des* phénomènes géographiques avec *un* SIG, sans préjuger non plus des outils exploités. La partie B de cette thèse présente les différentes étapes du développement de cette méthode :

– dans un premier temps, nous nous intéressons à l'étape de modélisation en information géographique qui prépare la constitution des données (chapitre B.1)

– puis nous revenons sur les différents moyens de représenter un phénomène géographique, pour aborder ensuite les différentes façons de gérer ces représentations dans un SIG. Nous proposons alors une méthode pour créer de nouvelles informations à partir de données dans un SIG) (chapitre B.2).

– enfin (chapitre B.3), nous montrons que les données construites avec la méthode proposée permettent à la fois :

- une observation fine des phénomènes, améliorant leur observation d'un point de vue statique,
- et un suivi automatique dans le temps, permettant l'observation d'évolutions détaillées de ces phénomènes.

Ce dernier chapitre montre également la complémentarité entre l'approche méthodologique que nous avons adoptée et l'approche technique. Il explique en effet comment améliorer encore les observations en associant les deux : comment intégrer les outils de gestion d'informations temporelles sophistiqués pour saisir des évolutions détaillées avec la méthode.

La dernière partie de la thèse présente la mise en œuvre de la méthode générique proposée sur trois cas particuliers.

A.3.3.3. Trois contextes applicatifs différents pour tester la méthode proposée

– L' urbanisation « pavillonnaire ».

Constatant l'intérêt suscité par les « zones d'habitat individuel » auprès de thématiciens issus de domaines variés, constatant également qu'il n'existe pas beaucoup de données pour décrire ces 'objets', nous avons souhaité mettre en place une méthode d'observation spécifique pour permettre de les suivre. Leur observation à partir de photographies aériennes mettant en évidence leur « texture typique » : petits *bâtiments* répartis de façon assez homogène, parfois de façon très régulière autour d'un *réseau routier* souvent structuré, nous avons alors pensé pouvoir qualifier les « *zones pavillonnaires* » à partir des « pavillons » qu'elle regroupe. Et plus généralement nous avons eu l'idée de décrire des évolutions d'une zone pavillonnaires à partir d'évolutions survenues sur le groupe de ses pavillons. Cette application à beaucoup contribué à l'élaboration de la méthode proposée, pour valider ses principes et tester sa faisabilité.

Pour illustrer nos propos nous évoquerons également deux autres cas d'application :

– L'imperméabilisation des sols

Le suivi de l'imperméabilisation des sols dans un SIG part du même constat que le phénomène d'urbanisation pavillonnaire : bien l'on puisse facilement identifier les surfaces imperméables à l'aide d'une photographie aérienne, il est plus difficile de pouvoir les observer dans un SIG puisqu'il n'existe pas de base de données régulièrement mise à jour pour fournir cette information. Cependant dans la cadre spécifique retenu pour cette application (les études d'aménagements hydrauliques), il est intéressant de noter que ce n'est pas tant l'information géographique 'zone imperméable' (avec sa composante géométrique, cf. p. 30) qui est souhaitée par les thématiciens, mais c'est l'information sémantique *superficie* de la *surface imperméabilisé* qui est recherchée, car elle permet alors de déterminer un indicateur essentiel pour les analyses : le *coefficient d'imperméabilisation* (donné par le ratio : surface imperméabilisée sur la surface totale). Ainsi, pour répondre au besoin des thématiciens de pouvoir disposer de la valeur de ce ratio pour toutes les zones susceptibles d'être ou de devenir une *zone d'études d'aménagements hydrauliques*, nous avons envisagé la possibilité de déduire la superficie de la zone imperméabilisée à partir d'une estimation de la surface imperméabilisée générée par chaque objet topographique élémentaire situé dans la zone d'étude. Pour tester le principe de cette méthode d'évaluation du *coefficient d'imperméabilisation* nous avons travaillé en particulier sur l'imperméabilisation des *zones industrielles*.

– L'ambiance urbaine

L'observation de l'ambiance urbaine telle qu'attendue dans le contexte thématique dans lequel nous nous sommes situés (la gestion de centre-ville) comportait des dimensions difficiles à intégrer a priori dans un SIG (comme l'esthétique, la convivialité). Cette observation a ainsi demandé un travail de formalisation technique important pour définir des éléments de caractérisation qui puisse être intégrés au SIG sous forme de données géographiques, mais déjà avant cela pour définir à partir de quel 'objet géographique' nous allions suivre ce phénomène.

Ces cas sont présentés dans la partie C et des éléments d'application sont proposés en Annexe 7.

B.1. Entre données et phénomène observé : la formalisation des informations géographiques

B.1.1. L'information géographique comme médiateur entre une approche thématique et une approche technique

L'utilisation de systèmes informatiques **(SIG) pour observer des phénomènes géographiques ne pose pas seulement un problème technique** (il ne s'agit pas uniquement de développer la capacité des outils SIG à gérer la dimension temporelle). Comme le montre la définition d'«objet » géographique rendue nécessaire dès l'introduction (cf. p. 22), ces recherches renvoient également à des problèmes plus cognitifs liés à la représentation, parmi lesquels les problèmes de vocabulaire ne sont qu'un symptôme [Nyerges et al, 1995]. Dans ce chapitre nous illustrerons ces problèmes sur plusieurs exemples. Nous montrerons l'intérêt d'introduire une étape de formalisation qui ne soit pas uniquement technique, ni uniquement thématique. Cette étape de formalisation nous servira en particulier à assurer une meilleure modélisation des phénomènes géographiques dans le SIG, où meilleure veut dire à la fois : qui convient pour l'observation à réaliser, qui soit capable d'être enrichie et que l'on peut exploiter commodément dans les SIG sans développement complexe. En nous appuyant ensuite sur cette modélisation, nous construirons une méthode pour enrichir la description d'un phénomène (chapitre B.2) avec des informations dont nous montrerons qu'elles sont également simples à suivre dans le temps (chapitre B.3)

B.1.1.1. Entre phénomène et données : l'information géographique

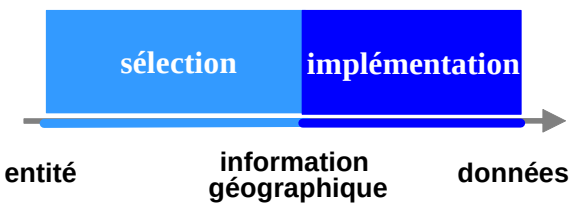


Figure B.1.1. *Trois niveaux conceptuels pour observer un phénomène avec des données dans un SIG*

Si dans un premier temps, on distingue de façon schématique trois niveaux conceptuels différents dans le processus de formalisation qui dans un SIG mène du phénomène géographique aux données (cf. figure B.1.1):

- le niveau du phénomène (celui sur lequel on veut acquérir des connaissances),
- le niveau des informations géographiques (celui sur lequel on raisonne et effectue des analyses),
- le niveau des données (celui des traitements, des mesures, des calculs),

l'information géographique apparaît alors comme un intermédiaire entre une approche thématique centrée sur le phénomène (ex. : *comment réduire le phénomène de congestion des voies de circulations ?*) et une approche technique des données géographiques (ex. : *comment représenter le 'trafic' sur le 'réseau routier' ?*) (fig. B.1.2). Or, le changement d'approche peut soulever des difficultés [Nyerges, 1995],

- soit parce que les SIG ne sont pas capable d'intégrer les informations du thématicien :

« When basic language in the tool domain are not sufficiently developed, then one can not express the problem in the terms of the tool language. As corollary, when the language in the tool is not sufficiently developed, then the tool prohibits a user from expressing what is known about the problem to carry out problem solving» [Nyerges, 1995, p.72].

C'est le cas par exemple avec le temps ; les SIG ont du mal à intégrer la dimension temporelle (cf. chapitre A.3). C'est également le cas avec les *objets* possédant des limites floues, car les SIG modélisent les objets à l'aide des primitives géométriques : polygone, polyligne et point, dont les limites sont strictement définies ([Girard, 2004], [Saint-Joan, 1996]).

- soit parce le problème géographique n'est pas formulé de façon suffisamment proche des modèles utilisés dans les SIG pour pouvoir être implémenté sous forme de données.

« When basic knowledge constructs in the problem domain are not sufficiently developed, then one does not know how to express problems in a particular setting » [Nyerges, 1995, p.72].

C'est le cas par exemple lorsque l'on veut observer le phénomène de l'ambiance urbaine, qui demande de saisir une information localisée à laquelle il est difficile d'associer une composante géométrique (cf. Partie C). Le cas de la «qualité de vie» implique de même un travail spécifique pour le définir en terme d'informations géographiques observables dans un SIG [Bararino, 2005].

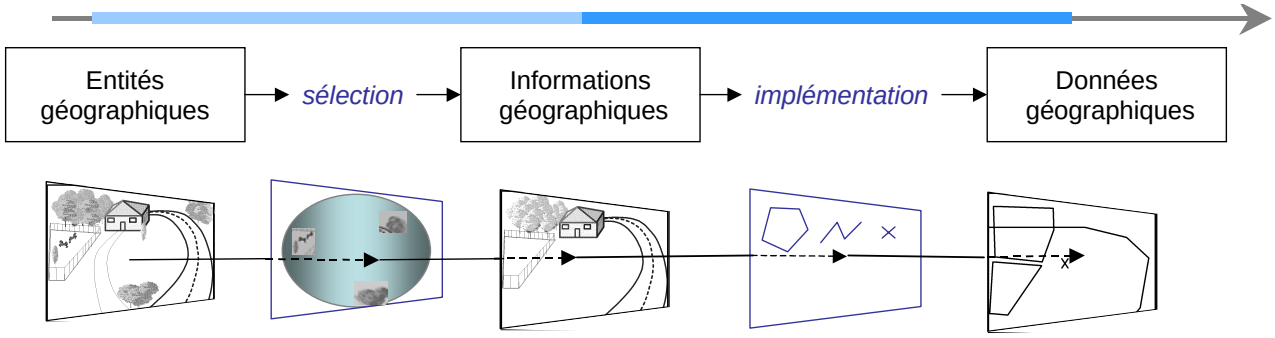


Figure B.1.2. *Deux processus pour passer des entités géographiques aux données : la sélection et l'implémentation*

B.1.1.2. Deux informations géographiques : celle du technicien, celle du thématicien

Le producteur de données et le thématicien, qui interviennent parfois de façon autonome, aboutissent à des conceptions différentes des informations géographiques, de part leurs compétences et leurs objectifs distincts (fig. B.1.3). Le chapitre A.3 a montré le problème d'interprétation des informations temporelles (ex. : la mise à jour technique peut introduire des changements qui ne correspondent pas à des évolutions réelles ; il peut y avoir des corrections d'erreurs, des mises à niveau de la base). La façon de spécifier les informations géographiques présente la même difficulté à s'entendre.

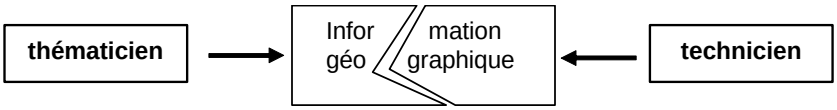


Figure B.1.3. *Deux approches différentes de l'information géographique*

Une approche technique pour spécifier les informations géographiques

Une façon de spécifier les informations géographiques consistent à les répartir en grands *thèmes génériques*, puis à préciser quels sont les 'objets' retenus dans chacun de ces thèmes (dans le document appelé *spécifications de contenu*). Cette démarche correspond à celle adoptée par les producteurs de données comme l'IGN par exemple pour élaborer la BDTOPO® (fig. B.1.4). Cette approche des spécifications est fondamentalement technique. En annexe 2, nous montrons que ces grands *thèmes* s'interprètent d'ailleurs comme le résultat d'aspects techniques, à la fois historiques, informatiques, mathématiques, et enfin, pratiques. Dans cette approche chaque entité est représentée par à un seul 'objet' ; il s'agit ici de représenter un territoire (plutôt a-thématique, topographique dans le cas de la BD TOPO®) selon un point de vue caractérisé par *l'échelle d'observation*[1]. Les données qui en résultent servent généralement de *support*, de *référentiel* à d'autres données[2]: elles constituent un base de départ pour les thématiciens pour élaborer leur propre base de données.

1. A l'époque où l'observation des informations géographiques s'effectuait à partir de représentations cartographiques papier, l'échelle d'observation correspondait à l'échelle de représentation. Avec la représentation numérique, il est possible avec les outils de zoom d'afficher la même base de données à différentes échelles. L'échelle d'observation correspond alors à la *« gamme d'échelle d'utilisation »* de la base.

2. Des *données support* (aussi appelées données de référence) sont des données existantes qui permettent à l'utilisateur de localiser ses informations pour pouvoir les exploiter dans un SIG.

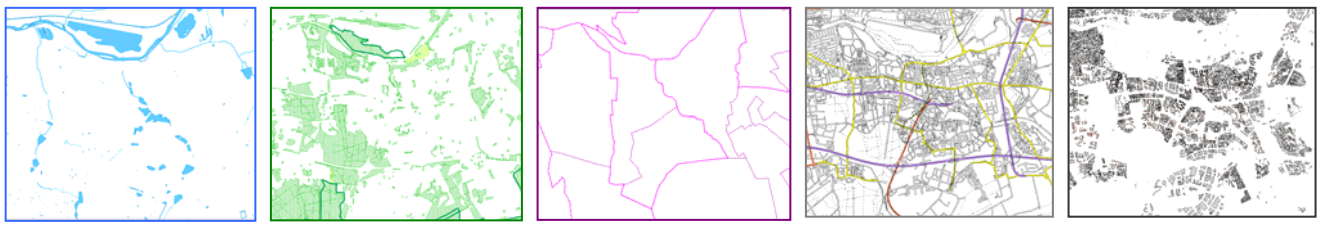


Figure B.1.4. *Les grands thèmes de la BD TOPO®*

Une approche thématique pour spécifier les informations géographiques

Les données supports jouent un rôle essentiel auprès des thématiciens, pour implémenter les informations géographiques qu'ils sélectionnent, pour formaliser les *objets géographiques* qu'ils spécifient. En effet, contrairement au cas précédent, certains outils-SIG réalisent une *sélection multiple* et proposent plusieurs représentations en *objets géographiques*

pour une même entité (fig. B.1.5) ; plusieurs points de vue sur une même entité géographique peuvent cohabiter et les données support peuvent aider à les intégrer (ex. : l'INSEE utilise les 'îlots' pour porter des attributs sur la population, les commerces, les entreprises, l'habitat, etc.). Le SIG thématique tente ainsi de répondre au besoin d'effectuer des réflexions croisées sur les problématiques[3] : par exemple, pour mettre en adéquation le PLU (Plan Local d'Urbanisme) avec les besoins en équipements, pour étudier les relations entre les accidents et la forme des tissus urbains, etc. Les organismes comme L'APUR ou L'IAURIF dont les SIG (données) offrent à l'observation des informations très variées, contribuent ainsi à l'étude de phénomènes combinant de nombreux aspects.

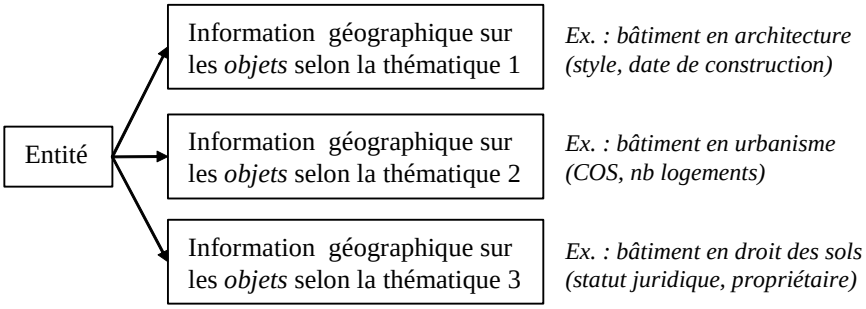

Entité
Information géographique sur les *objets* selon la thématique 1
Information géographique sur les *objets* selon la thématique 2
Information géographique sur les *objets* selon la thématique 3
Ex. : bâtiment en architecture (style, date de construction)
Ex. : bâtiment en urbanisme (COS, nb logements)
Ex. : bâtiment en droit des sols (statut juridique, propriétaire)

Figure B.1.5. *Plusieurs représentations pour une même entité géographique*

La différence pouvant exister entre les représentations en informations géographiques thématiques (ou *objet géographique*) n'est pas la différence évoquée en début de paragraphe entre *l'information géographique thématique* et *l'information géographique technique*. La différence entre *objets géographiques* est le résultat de la différence de point de vue sur les entités géographiques entre thématiciens (différence de *systèmes de référence thématique* cf. A.2.2.1).

Si l'un d'entre eux choisit d'intégrer cette différence dans la représentation qu'il a de l'entité, alors la différence disparaît : il n'y a plus qu'*un objet géographique* caractérisé de façon plus riche. Par exemple, un architecte peut choisir d'ajouter aux qualités architecturales d'un *bâtiment*, sa fonction administrative et des caractéristiques relatives aux réglementations d'urbanisme. Il travaillera alors sur une représentation de l'entité bâtiment plus complète.

L'utilisateur peut aussi choisir d'ignorer les spécificités thématiques de l'entité autres que celles observées par lui. Alors la différence n'existe plus ; elle n'est pas.

En revanche, il existe bien une différence entre l'information géographique telle que sélectionnée par le thématicien et celle implémentée dans les données SIG. Là où l'utilisateur ne voit qu'un seul *objet géographique* qu'il aimerait étudier en tant que tel, la représentation dans le SIG lui impose parfois de différentier plusieurs 'objets géographiques' informatiques pour faciliter la gestion et les traitements.

EXEMPLE 1 : Pour une *route*, la représentation sous forme de polyligne est requise pour calculer des itinéraires, car les fonctionnalités topologiques utilisées pour ce type d'analyse se fondent sur la théorie des graphes et travaillent donc à partir de lignes. En revanche pour calculer précisément une surface (par exemple dans le cadre de la gestion de l'entretien de la route), il est nécessaire de disposer d'une représentation surfacique pour prendre en compte les variations de largeur de voie le long de l'axe) (fig. B.1.6).

3. On évoque alors le *SIG fédérateur*, le *SIG pluridisciplinaire* ou *interdisciplinaire* (*«L'interdisciplinarité favorise la fertilisation croisée »* [Zaninetti, 2005]) et peut-être même *transdisciplinaire* !

Selon [Nicolescu, 1996] :

Interdisciplinaire concerne le transfert des méthodes d'une discipline à l'autre… mais dont les finalités restent inscrites dans la recherche disciplinaire,

Pluridisciplinaire concerne l'étude d'un objet d'une seule et même discipline par plusieurs disciplines à la fois,

Transdisciplinaire est à la fois entre les disciplines, à travers les différentes disciplines et au-delà de toute discipline. Elle a comme finalité la compréhension du monde présent dont un des impératifs est l'unité de la connaissance.

Nous noterons ici que *l'interdisciplinarité* telle qu'évoquée par [Zaninetti, 2005, avant propos] ne semble être conçue qu'en *terme thématique*. En effet, il ne cite que des exemples de sciences et de disciplines thématiques, mais n'y adjoint pas de disciplines techniques, qui pourtant contribuent aux analyses dans les SIG.
Or l'enjeu de l'observation avec un SIG ici est bien d'intégrer des approches différentes, pas seulement thématiques, mais surtout thématiques *et* techniques.

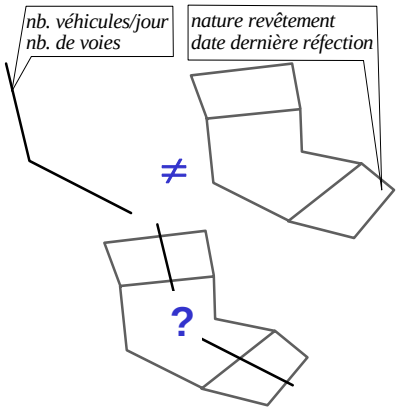


Figure B.1.6. *Une même entité est parfois représenté de plusieurs façons dans le SIG : l'exemple d'une route*

EXEMPLE 2 : Pour observer un même *objet* à des échelles différentes avec un SIG, l'utilisateur est obligé de travailler indépendamment sur des représentations différentes. En effet, suivant les principes cartographiques, la représentation d'un *objet géographique* doit être fonction de l'échelle d'observation. Or, dans une SIG, la gestion automatique et l'analyse conjointe de représentations différentes pour un même objet est un problème technique complexe dit : *problème de la gestion de représentations multiples* (cf. paragraphe B.2.2.1.1).

Ainsi, producteur et utilisateur de données géographiques, et plus généralement technicien et thématicien de l'information (géographique) se trouvent dans une même chaîne de construction d'informations, chacun à un bout du processus d'observation d'un phénomène à partir de données : les données construites par les techniciens servent aux études menées par les thématiciens. Cette relation d'enchaînement les situent dans une relation de dépendance : les analyses du thématicien dépendent de travaux qui ne sont pas les siens. Or, si leurs préoccupations sont proches, leur façon de concevoir l'information géographique peuvent être fondamentalement différentes, en raison de leur approche respective.

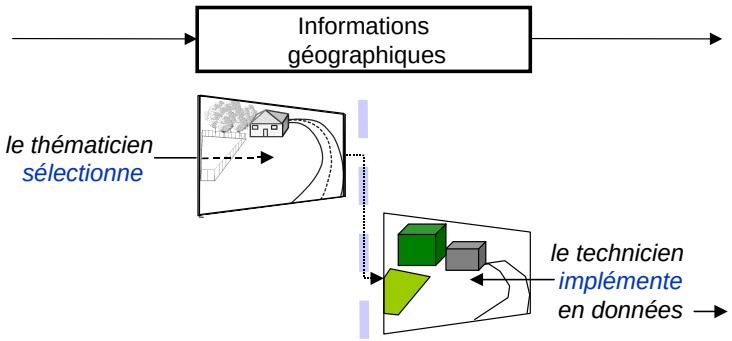


Figure B.1.7. *Entre phénomène et données : différentes informations géographiques*

B.1.1.3. *L'entre-deux*, entre technique et thématique

B.1.1.3.1. Introduction à l'entre-deux

En soulignant cette différence, notre objectif n'est pas d'opposer le point de vue technique et le point de vue applicatif, sous prétexte d'*objets* distincts (fig. B.1.7). Pour nous, il s'agit d'y voir une différence (une relation d'altérité) et non un différend (désaccord résultant d'une différence d'opinion, d'une opposition d'intérêts entre deux ou plusieurs personnes). Et même le concept de différence s'avère encore trop fort.

« Le trait de la différence marque un bord, celui d'un ensemble opposé à un autre, ou opposé au reste, rassemblant les « gens » du même bord par différence avec les autres » [Sibony, 1991, p.14].

Il ne s'agit pas pour nous de confiner chacun à une fonction spécialisée dans une vision morcelée du processus de construction des connaissances, avec d'un côté ceux qui auraient en charge la définition du *monde représenté* et de l'autre ceux qui auraient pour mission de définir le *monde représentant* (pour reprendre les termes posés par la théorie générale de la représentation de Palmer citée par [Hernandez, 1995, p. 49 (45-59)]), ce que l'on ferait en évoquant la notion « géographique » de *frontière* ou de cloisonnement. La notion *d'interface* est plus riche, car elle emprunte au domaine informatique l'idée de transfert d'informations et à la chimie une certaine épaisseur de la limite commune, mais elle ne nous paraît pas convenir non plus. Nous lui préférons celle *d'entre-deux*, proposée par

[Sibony, 1991] (fig. B.1.8) que nous présentons de façon plus détaillée en Annexe 3.

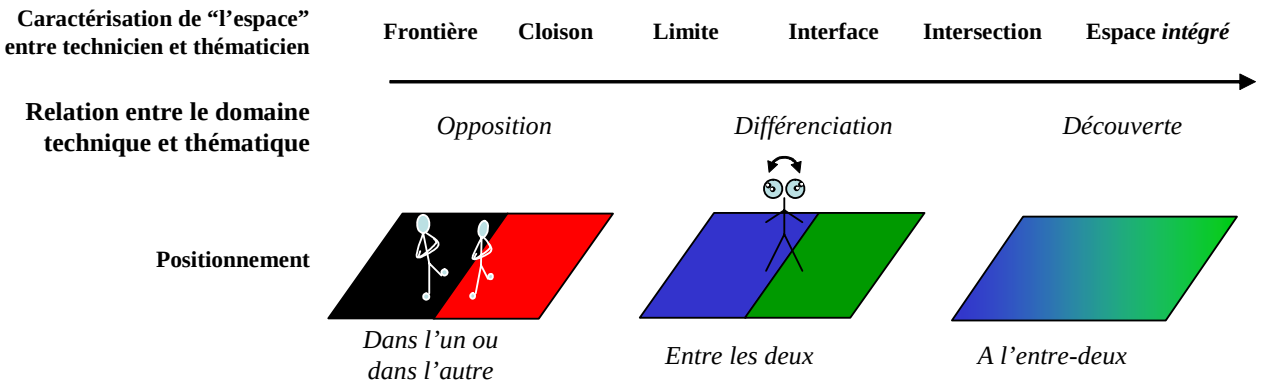


Figure B.1.8. *Comment aborder l'information géographique partagée entre domaine thématique et domaine technique ?*

Il nous semble en effet qu'il existe entre le producteur de données et l'utilisateur thématicien, un espace à part entière où les questions ne se posent pas en terme de définition de LA bonne représentation, mais où les questions s'appuient sur les connaissances et les points de vue de chaque approche pour les intégrer, où le potentiel de chaque domaine de connaissances peut être exploité au mieux, pour élaborer un SIG permettant d'observer un phénomène. *« Having knowledge in the problem domain or knowledge in the tool domain does not ensure that one is sufficiently skilled to carry out GIS data processing. Skill arises not only from a mutual development of problem domain knowledge and tool domain knowledge, but from an ability to carry out acts as a synthesis of these two kinds of knowledge »* [Nyerges, 1995]. C'est à cet endroit précis que nous avons voulu nous positionner pour aborder le problème de l'observation des phénomènes à l'aide de données géographiques, dans une démarche qui considère l'information géographique comme un *médiateur* entre l'approche thématique et l'approche technique (fig. B.1.9).

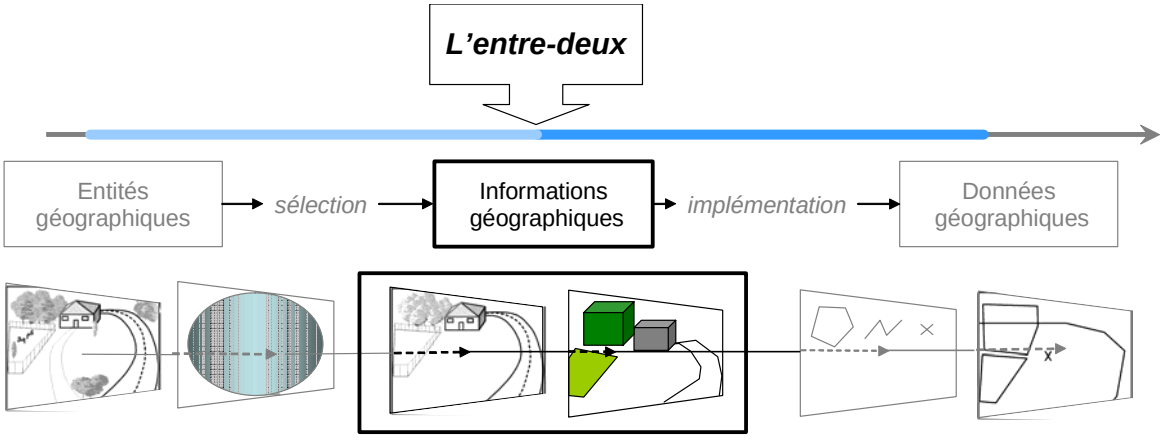


Figure B.1.9. *Des informations géographiques situées à l'*entre–deux *pour observer des phénomènes géographiques dans un SIG*

B.1.1.3.2. L'analyse des informations géographiques à l'entre-deux

L'adoption de *l'entre-deux* comme point d'observation représente un changement de perspective[4]. De façon générale, ce positionnement permet d'expérimenter des conceptions différentes, en particulier sur des termes pouvant être par ailleurs fondamentaux. Lorsque les conceptions ne sont pas contradictoires, établir un lien entre elles permet de profiter de l'apport

4. De façon spécifique, elle nous a aidé à éviter une conception uniquement informatique.

des différents domaines dans lesquels ces conceptions opèrent : soit pour disposer d'un domaine de conception « fédérateur » plus large (fig. B.1.10a) (comme pour la définition du terme « SIG », cf. note p. 21), soit pour disposer de compétences supplémentaires pour mettre en œuvre un processus d'analyse complexe dont les étapes s'inscrivent dans des domaines différents (fig. B.1.10b) (voir les différents sens pris par « l'objet géographique »). Le double terme « analyse spatiale » considéré à *l'entre-deux*, nous semble procéder des deux cas.

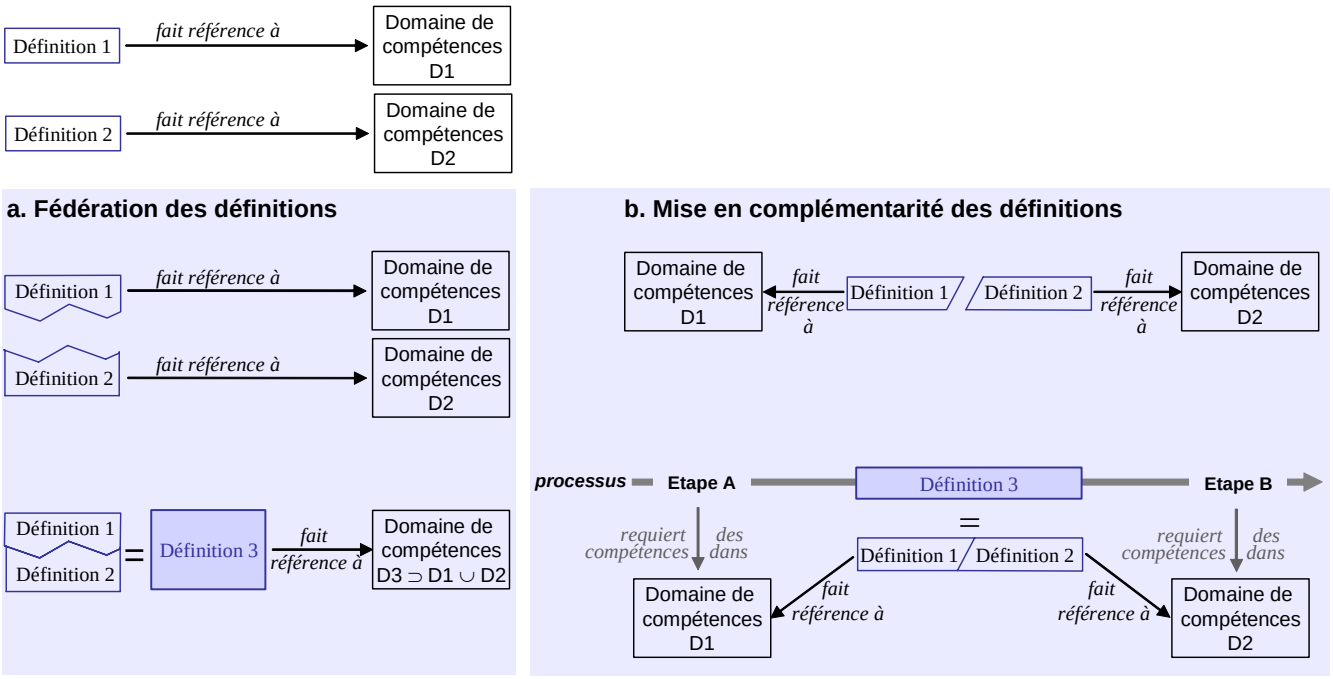


Figure B.1.10. *Définition à l'entre-deux d'un terme conçu par ailleurs de façons variées*

Approche technique

Dans le domaine technique des logiciels SIG, le terme « d'analyse spatiale » est souvent employé pour désigner un traitement élaboré, sur des informations géographiques, par opposition à une requête simple de type Où ? ou Quoi ?. Il peut s'agir de traitements sur :

- la composante sémantique (plutôt statistiques),
- la composante géométrique (études morphologiques),
- la composante topologique (ex. : analyses de graphes ou de partitions),

voire des traitements complexes faisant intervenir plusieurs composantes.

«A*nalyse spatiale : les opérateurs topologiques.*
L'analyse spatiale consiste principalement à déterminer les interactions spatiales entre les objets géographiques et à partir de ces interactions, dériver de nouvelles informations. Elle repose donc essentiellement sur la détermination d'intersections, d'inclusions, de proximités. » [Denègre et Salgé, 1996, p.81],

Ces différentes analyses font appel à des expertises mathématiques différentes, à des outils et méthodes spécifiques. Le niveau de compétence requis pour maîtriser ces analyses tend à les cloisonner et ainsi à spécialiser les définitions[5]. La définition du terme double « analyse-spatiale » est floue également parce qu'il n'a pas de définition propre dans le domaine technique. Sa définition est déduite de l'association des deux termes : « analyse » et « spatiale ». Or, il existe plusieurs analyses : {topologique/ morphologique / statistique}, {statique / cinématique / dynamique}, {dis-

5. Dans une expérience menée au sein du laboratoire COGIT le 19/1/2005 à l'IGN, nous avons posé la question : « qu'est ce que l'analyse spatiale pour vous ?». Les réponses ont mis en avant une grande variété de définitions : certaines étaient centrées autour de l'analyse de la composante géométrique ou de la topologie, d'autres étaient plus proches des statistiques, certaines étaient techniques, d'autres plus thématiques…

crète / continue} ; et il existe plusieurs formes de spatialisation (ce qui nous conduira d'ailleurs à développer une typologie des informations géographiques, cf. B.1.2.2.2) [Bordin, 2005a].

Approche thématique

Le terme « d'analyse spatiale » se retrouve également dans le domaine thématique. Grâce à la géographie, il fait l'objet d'une définition précise et formalisée [Pumain et Saint Julien, 1997]. Les dictionnaires de géographie proposent ainsi une définition propre à ce double terme :

«Ensemble de méthodes mathématiques et statistiques visant à préciser la nature, la qualité, la quantité attachées aux lieux et aux relations qu'ils entretiennent – l'ensemble constituant l'espace-, en étudiant simultanément attributs et localisations » [Brunet, 1992], ou encore

« L'analyse spatiale met en évidence des structures et des formes d'organisation spatiale récurrentes, que résument par exemple les modèles centre-périphérie, les champ d'interaction de type gravitaire, les trames urbaines hiérarchisées, les divers types de réseaux ou de territoires, etc. Elle analyse des processus qui sont à l'origine de ces structures, à travers des concepts comme ceux de distance, d'interaction spatiale, de portée spatiale, de polarisation, de centralité, de stratégie ou choix spatial, de territorialité... Des lois de la spatialité relient ces formes et ces processus, et sont intégrées dans des théories et des modèles du fonctionnement et de l'évolution des systèmes spatiaux…» [Hypergéo.free.fr]

Cette conception est centrée autour des modèles, outils et concepts géographiques, mais elle s'intéresse aussi à la mise en œuvre technique des analyses. Ainsi, la définition de D. Pumain dans [Hypergéo.free.fr] mentionne la variété des moyens : *« Les méthodes mobilisées par l'analyse spatiale sont très diverses », «… S'appuyant sur des méthodes statistiques et des modèles mathématiques, utilisant les cartes, les systèmes d'information géographique (S.I.G.) et divers outils de simulation, tout en intégrant les résultats d'enquêtes sur les comportements dans l'espace et ses représentations… »*. Elle évoque également le statut privilégié des analyses statistiques : *« L'emploi de la statistique classique conduit à oublier dans un premier temps la localisation des objets et, complété par la cartographie, retrouve ensuite, dans une analyse en termes de résidus, les effets spatiaux. Au contraire, les méthodes de la géostatistique et de la morphologie mathématique, regroupées dans la statistique spatiale, analysent directement les informations géocodées »*.

B.1.1.3.3. Analyse à référence spatiale

Dans le cadre de ces recherches qui établissent un lien entre l'observation de phénomènes géographiques et l'étude de données dans un SIG, nous proposons de considérer l'« analyse spatiale » à l'*entre-deux* pour tenter de :

- fédérer les différentes approches techniques (fédération des définitions cf. figure B.1.10a), pour inciter à utiliser une gamme plus large d'analyses et ainsi à mieux exploiter le potentiel d'informations contenu dans les données,

- intégrer les différentes approches techniques à l'analyse thématique pour mieux traiter le problème de l'observation de phénomène *avec* un SIG (mise en complémentarité des deux approches technique et thématique cf. figure B.1.10b).

Cette conception de « l'analyse spatiale » a contribué au développement de la méthode proposée pour observer des phénomènes géographiques avec un SIG. En effet, elle permet de construire une caractérisation plus riche des 'objets', pour une observation plus détaillée. Elle permet également d'envisager une observation multiniveaux dans les SIG, alors que le problème technique des représentations multiples semble s'y opposer. En revanche, si le positionnement choisi pour ces recherches est *l'entre-deux*, la méthode proposée, quant à elle, se situe résolument dans le domaine technique ; il s'agit de travailler sur la modélisation des 'objets géographiques' dans les SIG pour une meilleure exploitation à la fois des fonctionnalités d'analyses et des informations contenues dans les composantes géométriques, sémantiques et topologiques. L'objectif final est de permettre aux utilisateurs de disposer d'informations nombreuses et facilement exploitables pour suivre des phénomènes géographiques.

La définition de « l'analyse spatiale » ou plutôt de l'« analyse à référence spatiale » (pour réserver le terme d'« analyse spatiale » à la géographie qui l'emploie déjà dans un sens précis) doit s'appuyer sur celle des géographes, par exemple celle de D. Pumain. En effet,

- cette définition existe et elle est opérationnelle : elle est la référence commune à de nombreux utilisateurs, contrairement à une « analyse spatiale » plus technique qui semble dépendre des domaines techniques dont sont issus les outils et méthodes employés.

- cette définition constitue un point d'entrée vers de nombreux domaines thématiques : *«… l'analyse spatiale est employée par bien d'autres disciplines que la géographie, en économie spatiale (ou science régionale), histoire, agronomie, archéologie, sciences de l'environnement, etc. »*

- elle est compatible avec les différentes « définitions techniques », puisqu'elle fait le lien avec les méthodes et outils utilisés .

- pour ces recherches, elle peut aider à mettre en complémentarité les conceptions, car elle présente à la fois l'aspect applicatif le « pourquoi ? » et l'aspect fonctionnel le « comment ? »,

- enfin, cet énoncé montre l'« *attention particulière apportée en analyse spatiale à la définition de l'échelon géographique considéré, du niveau d'observation, qu'il s'agisse du niveau " microscopique " des acteurs individuels ou d'agrégats spatiaux définis à des niveaux meso- ou macro-géographiques* ». Or le *niveau d'observation* est une notion importante dans cette thèse puisque qu'il nous servira à articuler l'observation des *objets géographiques* avec la production de données dans le SIG (cf. chapitre B.2).

En revanche, notre définition de l'« analyse à référence spatiale» fait une place plus importante aux méthodes et outils d'analyses techniques (le comment ?) ; en particulier elle intègre de façon plus marquée les analyses morphologiques et topologiques. Il importe en effet pour nous de valoriser ces analyses, car elles sont un apport important des SIG[6] ; poser l'utilisation des SIG implique pour ces travaux de tirer profit de ces outils ET de surmonter les contraintes engendrée par leur utilisation. Par ailleurs, des travaux de recherche ont montré la capacité des analyses morphologiques et topologiques à construire des indicateurs utiles à la caractérisation des informations géographiques [Duchêne, 2003], [Gleyze, 2005], [Claramunt, 2005] (fig. B.1.11). Ces indicateurs peuvent aider à caractériser les 'objets géographiques' comme le 'parcellaire', le 'bâti', l''emprise d'une ville' ou d'une agglomération, les 'réseaux', etc. [Gleyze, 2005] situe d'ailleurs explicitement les indicateurs mathématiques qu'il élabore pour décrire des graphes (des réseaux), dans la perspective d'analyses

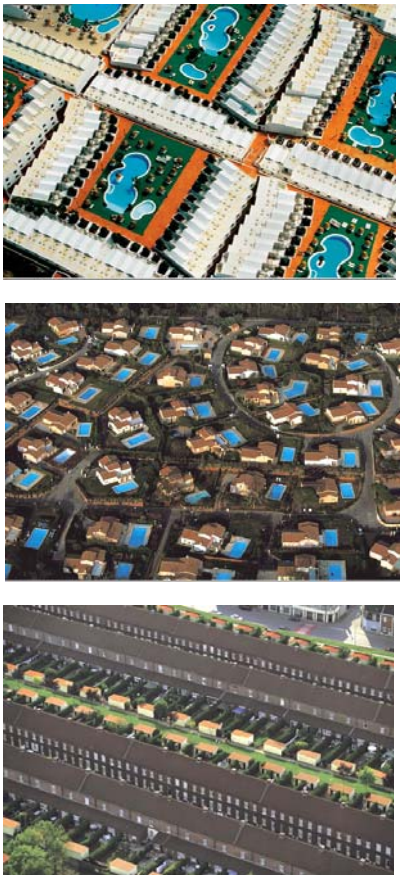


Figure B.1.11. *La morphologie et la topologie pour caractériser les objets géographiques (Photographies de A. Bertrand : www.yannarthusbertrand.com)*

6. Depuis l'origine, les SIG sont les outils qui gèrent et étudient conjointement les trois composantes de l'information géographique. Par rapport à la cartographie numérique, ils exploitent la compo-sante géométrique comme une représen-tation de l'emprise des objets et pas seulement comme un support de variables graphiques (ex. : couleur, taille, forme). Ainsi, mesurer la superficie du polygone représentant une école a un sens géogra-phique (c'est l'emprise au sol du bâtiment), tandis que mesurer la taille du symbole n'en a pas. L'étude de la composante géométrique apporte des éléments de connaissance sur l'objet géographique, complémentaires à celle de la composante sémantique (ex. : sur la régularité des formes d'un bâti, sur l'orthogonalité d'un maillage routier, sur l'homogénéité d'un parcellaire). Les SIG distinguent la représentation au sens de modélisation et la représentation au sens cartographique.

applicatives : « *Analyse structurelle des réseaux dans un contexte de risque* ». Néanmoins, ces analyses restent encore sous valorisées, malgré leur potentiel, car les outils et méthodes pour les mettre en œuvre ne sont pas encore suffisamment développés en standard dans les logiciels. Ils font l'objet de développements spécifiques, parfois dans le cadre de recherches techniques[7]. De fait leur visibilité se trouve restreinte au risque parfois d'être oubliées en tant qu'« analyse spatiale » ou d'être cantonné à être des analyses à référence spatiale *uniquement* technique. Pourtant, des mises en œuvre simples peuvent être effectuées (cf. B.2, p. 156) qui peuvent participer à des analyses plus thématiques. Le cas des *zones pavillonnaires* nous sert en partie pour l'illustrer.

7. ex. : les recherches en généralisation qui déduisent d'une base de données des représentations à des échelles plus petites que celles associées à la base.

B.1.2. A l'*entre-deux* : la formalisation technique

Face à la différence de point de vue entre technicien et thématicien, pour répondre au besoin de suivi temporel de phénomènes géographiques avec un SIG, on aurait pu :

– constater cette différence en concluant que les difficultés évoquées par [Nyerges, 1995] sont réelles et proposer de les surmonter pour un cas de phénomène particulier en construisant des données spécifiques pour l'observation de ce phénomène.

– adopter une approche uniquement technique et construire un modèle général qui améliore de façon globale la prise en compte de la dimension temporelle dans les SIG [Badard, 2000] et laisser ensuite au thématicien le soin de l'adapter à l'outil qu'il utilise, puis celui d'instancier ce modèle en fonction du phénomène étudié.

– confier au thématicien le soin d'optimiser lui même les spécifications de la bases de données qu'il souhaite observer, grâce à l'exploitation des moyens de gestion temporelle déjà disponibles dans les SIG (nous avons vu que bien intégrés les outils et méthodes existants permettent de construire des applications de suivi sophistiquées cf. A.3.3.1) et laisser au soin du technicien uniquement la production de la base de données ainsi spécifiée.

Mais ces solutions ne répondent pas au problème tel que nous l'avons posé dans cette thèse qui choisit de fixer a priori ni le phénomène, ni les outils, ni les compétences techniques. Dans ces recherches, il s'agit de concilier les deux pôles de l'observation.

- le pôle **« thématique »** où l'observation porte sur des entités pour étudier un phénomène

- le pôle **« technique »** où l'observation porte sur des données et tente d'en extraire des informations utiles à l'étude des phénomènes géographiques.

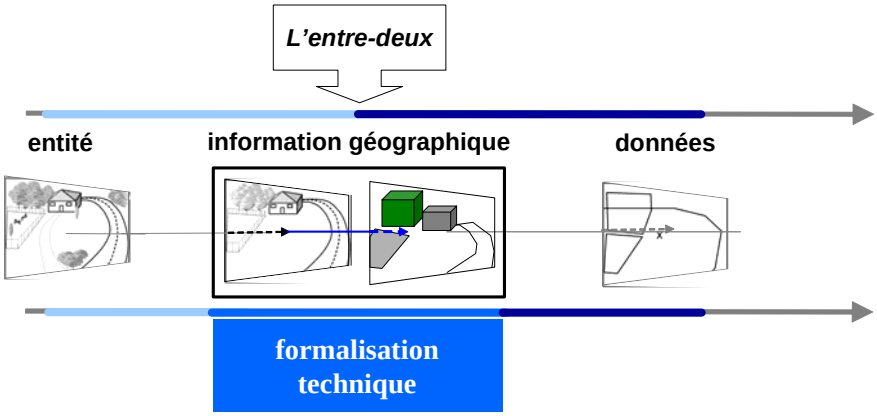


Figure B.1.12. *La formalisation technique comme étape intermédiaire entre la formalisation conceptuelle et l'implémentation en données*

Pour cela, nous proposons d'introduire à *l'entre-deux* une étape de formalisation intermédiaire entre la formalisation conceptuelle thématique qui sélectionne parmi les entités les *objets géographiques* à observer et leurs caractéristiques et l'étape d'implémentation en données géographiques dans un SIG (fig. B.1.12).

B.1.2.1. L'étape de formalisation technique

B.1.2.1.1. La formalisation technique avant l'implémentation en données

L'*implémentation* en données est un processus *technique*. Il s'agit d'intégrer les informations géographiques dans le SIG, sous forme de données, en spécifiant la forme des différentes composantes : géométrique, sémantique et topologique. (Par exemple, la 'route' est représentée par son axe, sous forme d'un linéaire, avec une précision de 1,5 m, décrite par les attributs : nombre de voies, largeur, sens de circulation. Les 'ponts' sont des 'tronçons de route' particuliers qui sur un graphe n'autorisent pas la connexion avec le 'tronçon de route' qu'ils surplombent). Techniquement, il importe de décomposer l'implémentation en deux étapes :

– une étape de *pré-implémentation*, qui correspond à ce que nous proposons d'appeler la *formalisation technique* (qui correspond à la *modélisation conceptuelle* en informatique)
– et l'étape *d'implémentation* proprement dite (qui correspond à la *modélisation logique* en informatique).

Pour ceux qui ont peu d'expériences pratiques des logiciels de SIG, les deux étapes peuvent sembler se confondre ; le travail pour spécifier les informations géographiques peut apparaître comme un simple problème de codage informatique. Pourtant, il s'agit bien de deux modélisations différentes :

– l'une, générique, est fonction de la modélisation dans les SIG ; elle produit des 'objets informatiques',
– l'autre, spécifique, dépend de la modélisation dans un logiciel-SIG particulier ; elle aboutit à des données.

L'étape d'*implémentation* proprement dite adapte la modélisation des informations aux spécificités du modèle du logiciel-SIG utilisé. En effet, bien que tous les SIG travaillent à partir d'une modélisation formelle standardisée (présentée dans le paragraphe A.1.1), chaque logiciel propose des variantes quant à son interprétation (ce qui soulève d'ailleurs des problèmes de format d'échange) :

EXEMPLE 1 : le concept *d'objet complexe* que l'on trouve dans le logiciel ArcInfo, mais pas dans les autres logiciels SIG, permet de définir un 'objet' comme un ensemble d'autres 'objets' : des objets simples (ex. : une 'route' à partir des 'tronçons de route' qui la constituent) (fig. B.1.13). L'objet complexe n'a pas de géométrie propre ; sa géométrie est déduite de celle des objets simples. En revanche, les attributs communs à tous les objets simples, relatifs à l'objet complexe (classement, n° de route) sont portés directement par l'objet complexe lui-même. Par exemple, les 'tronçons' ne sont décrits que par les attributs qui leur sont spécifiques (ex. : nombre de voies, largeur de voie). Ainsi, les attributs de l'objet complexe ne sont pas dupliqués ; ce qui techniquement garantit une plus grande qualité des données et assure à l''objet' géographique complexe une existence propre : on peut le sélectionner directement et enrichir sa description.

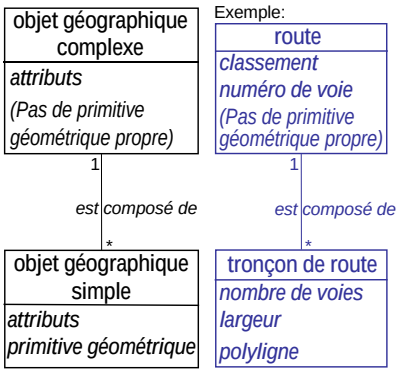


Figure B.1.13. *Concept* d'objet complexe *application sur l'exemple de la route*

EXEMPLE 2 : Le modèle du logiciel Géoconcept propose un autre concept le *type/sous-type* ; par exemple on peut y définir un type 'route' et ses sous-types : 'route départementale', 'route nationale', 'autoroute'. On peut ainsi travailler sur l'objet 'réseau routier' et associer à cet 'objet' des attributs commun à toutes les 'routes'. Ces attributs peuvent être des attributs techniques (ex. : date de production, nom du collecteur) ou des attributs thématiques (ex. : type de revêtement, service responsable). En revanche, comme le modèle du logiciel Géoconcept ne propose pas d'objet complexe, le n° de route sera ici associé à chaque 'tronçon'.

Dans [Bordin 2002, p.127], nous montrons les avantages à distinguer les deux étapes de formalisation. Cette distinction permet d'éviter en particulier
- une sous exploitation des capacités du logiciel,
- des difficultés dans les échanges avec les partenaires,
- un manque de portabilité des données et des applications, d'un logiciel à un autre.

Dans ce qui suit, nous ne prendrons pas en compte les variations de modèles propres aux logiciels-SIG. Nous considérons que l'étape d'implémentation fait partie de la mise en œuvre spécifique de la méthode générique et nous posons a priori que le technicien optimisera cette étape qui est de son ressort. De même, nous ne travaillerons pas sur l'étape de formalisation thématique au cours de laquelle l'utilisateur détermine les *objets* (et leurs caractéristiques) pertinents pour l'observation qu'il souhaite mener. Nous posons cette étape d'analyse du registre du thématicien comme un pré-requis à l'application de la méthode proposée (fig. B.1.14).

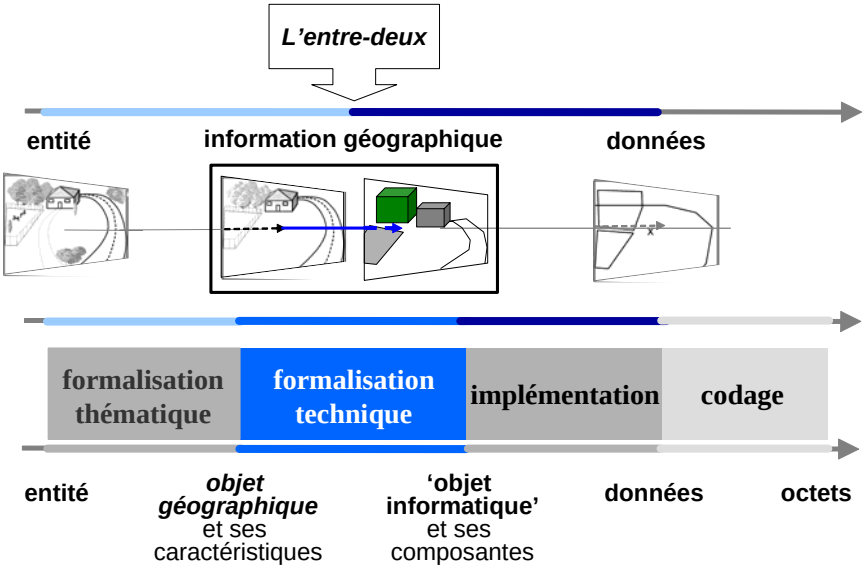


Figure B.1.14. *La formalisation technique comme préalable à l'implémentation*

En revanche, nous supposons que thématicien et technicien participent à la formalisation technique pour assurer que les *objets* (et leurs caractéristiques) identifiés par le thématicien sont « bien » modélisés en 'objets', c'est à dire de façon à ce que l'implémentation ultérieure sous forme de données permette de réaliser les analyses dont le thématicien a besoin. En effet, la façon de modéliser les *objets* en 'objets' influe sur les analyses que l'on peut effectuer. Nous l'avons évoqué avec l'exemple de la route (p. 94) en montrant que le choix de la composante géométrique devait dépendre de l'application (polyligne pour un calcul d'itinéraire, polygone pour un calcul de surface). De la forme du codage des attributs dépend également la capacité à réaliser certains traitements (par exemple si l'on code la 'date' sous forme de chaîne de caractère, il sera difficile d'ordonner les 'objets' par ordre chronologique). Le cas de l'ambiance urbaine nous sert à illustrer ces difficultés spécifiques, où la collaboration entre thématicien et technicien est indispensable pour savoir ce qui est recherché et choisir un codage adéquat : la formalisation des heures d'ouverture des 'magasins' a demandé un travail particulier pour permettre une analyse ultérieure de type time-géographie, de même la codification des 'types de rideau de fermeture' pour observer des aspects d'esthétique (cf. Partie C).

Mais la formalisation technique ne porte pas seulement sur la configuration des éléments participant à chacune des composantes. Dans ce qui suit nous montrons, qu'avant même de préciser ces aspects de configuration, il s'agit de réfléchir à l'élaboration de l'*objet* sous forme d' 'objet géographique', à la forme de ses composantes, car il existe plusieurs formes d'information géographique et toutes ne permettent pas de réaliser les mêmes analyses à référence spatiale.

B.1.2.1.2. La formalisation technique après la formalisation thématique

La formalisation technique ne se confond pas avec l'étape de formalisation thématique conceptuelle. Dans cette dernière le thématicien spécifie les *objets géographiques* et ses caractéristiques en fonction des outils et modèles qu'il veut utiliser. Et de fait cette modélisation peut parfois prendre des formes éloignées de celle attendue dans les SIG.

EXEMPLE 1 : L'utilisateur peut envisager une modélisation de la *pluie* sous forme de gouttes ponctuelles tombant de façon aléatoire sur un territoire. Mais, les SIG ne pouvant intégrer de ponctuels aléatoires, pour observer la pluie, il faut établir une autre modélisation et en particulier choisir une autre forme de localisation. On peut alors décider de représenter la 'pluie' par une estimation du nombre de mm d'eau tombée en moyenne sur un *territoire* et associer cette valeur sous forme d'attribut au polygone qui sert à représenter ce territoire. On peut également vouloir une définition plus précise et plus proche de la méthode de collecte de l'information. La 'pluie' peut alors être intégrée sous la forme du nombre de mm d'eau effectivement collecté par les appareils de mesure d'une station météorologique. Elle deviendra alors un attribut du point servant à représenter la *station météorologique*. Pour calculer la 'pluie' en d'autres points, on pourra alors utiliser des méthodes de krigeage. Notons que dans un cas, il faut disposer au préalable d'un 'objet-territoire' et dans l'autre d'un 'objet-station météo'.

EXEMPLE 2 : La modélisation en trois dimensions est nécessaire pour certains phénomènes (par exemple en géomorphologie). Elle est utilisée pour représenter beaucoup d'objets géographiques (ex. : le bâti, les carrefours complexes). Or les SIG ne proposent que des primitives 2D (polygone, polyligne, point), ce qui les empêchent de générer une vraie

3D [Abbott, 1998]. L'utilisateur doit donc adapter son modèle. Suivant les cas, il choisira ou bien

- d'exploiter les moyens de modélisation offerts par les SIG. Par exemple, pour des carrefours complexes, où plusieurs niveaux se superposent, il peut choisir d'utiliser la composante sémantique pour traduire la composition du carrefour (ex. : création d'un ponctuel 'carrefour complexe' avec comme attribut l'identifiant et l'altitude des 'tronçons' « superposés »). Cette option permet de gérer sous forme descriptive les informations nécessaires à l'analyse
- de s'adapter aux modèles de « simili 3D » proposés parfois sous forme de modules dans les SIG,
- d'exporter les analyses 3D dans des outils mathématiques dédiés et intégrer une partie des résultats dans le SIG (cf . certains travaux en hydrologie)
- de compléter les primitives standards par des développements supplémentaires de l'ordre de la recherche (par exemple, les travaux de [Poupeau et Bonin, 2006] qui proposent d'abstraire les *objets* à l'aide de primitives empruntées à la cristallographie).

Ces écarts de modélisation importants entre les *objets géographiques* tels qu'étudiés par le thématicien et les 'objets informatiques' tels que gérés dans les SIG nécessitent un travail d'adaptation à façon. Il ne peut être question ici de méthodologie générique pour l'utilisation des SIG. En effet, il s'agit de limiter ce à quoi doit renoncer l'utilisateur et de construire une solution qui lui convienne (ce qui de fait implique des traitements spécifiques) parmi la gamme des solutions techniquement possibles, cette gamme pouvant comprendre des solutions hors SIG (comme le montre l'exemple 2).

En revanche, il existe d'autres écarts de modélisation entre l'*objet géographique* tel que le thématicien souhaite le suivre et l''objet informatique' tel que formalisé dans le SIG. Ces écarts concernent uniquement la modélisation dans le SIG, la forme prise par la modélisation en 'objet géographique'. C'est dans cet écart que se situe exactement l'*entre-deux* entre thématicien et technicien.

B.1.2.2. Différenciation des informations géographiques

L'observation des *objets géographiques* dépend de façon essentielle de la modélisation en 'objets informatiques'. Et donc en retour, la modélisation en 'objets informatiques' dépendra de façon fondamentale de ce que l'utilisateur veut observer *et* des analyses qu'il veut effectuer.

Variété des informations géographiques

Thématiquement il existe une grande variété d'informations géographiques[8], 85% des informations gérées dans les divers systèmes d'informations sont potentiellement géographiques [Didier,1992]. Mais elles n'ont pas toute le même rôle, ni la même fonction pour décrire un territoire. Certaines sont principalement morphologiques, elles structurent l'espace géographique en réseaux (ex.: réseau routier, réseau d'assainissement, de télécommunication) ou en partitions (ex.: occupation des sols, parcellaire, unités administratives). Dans un SIG, elles possèdent généralement une composante géométrique sous forme de polyligne ou de polygone et sont souvent étudiées pour leurs relations topologiques avec les autres objets. D'autres informations, plus sémantiques[9], renseignent essentiellement sur des caractéristiques d'*objets géographiques*, comme les informations sur une entreprise (ex. : chiffre d'affaire, nombre d'employés, secteur d'activité) ou celles sur la population issues du recensement (ex.:

8. par exemple : une route décrite par le tracé de son axe et/ou par des caractéristiques sur le trafic, une étendue d'eau, une limite administrative, un point géodésique, une inondation, un plus court chemin entre un domicile et un travail, un client caractérisé par son adresse et par des attributs socio-démographiques, un bien immobilier, un crime, une épidémie, un ticket de caisse, une tombe dans un cimetière ou encore une ambiance urbaine… (Il ne manque qu'un raton laveur, et pourquoi pas? Les baleines sont bien information géographique lorsqu'on suit leurs déplacements dans un SIG après leur avoir implanté un émetteur GPS).

9. *sémantique* étant à entendre au sens géomatique de « qui porte sur des traits caractéristiques, sur des attributs, sur des descriptions non géométriques,… »

nombre d'habitants, de femmes/hommes, d'enfants, de retraités). Ces informations sont localisées simplement sous la forme d'un attribut d'une autre information (ex. : les informations sur la population d'une commune sont attachées à l'information géographique 'commune', les résultats d'analyses bactériologiques participent à la composante sémantique de l'information géographique 'tronçon hydrographique') ou bien sous la forme d'un point, grâce à un référent géographique (de type code postal ou Insee) et à une opération de géocodage[10] (comme dans le cas de l'entreprise *géocodée* à la ville, d'un client *géocodé* à son adresse).

Par ailleurs, une même entité géographique peut être observée et étudiée de façons très différentes. La formalisation dans le SIG doit restituer l'*objet géographique* de façon adaptée à son observation. Par exemple, pour suivre les évolutions de nature d'une *parcelle*, on peut définir un attribut : 'type d'occupation des sols' pour un objet 'parcelle'. En revanche, ce choix convient moins bien pour suivre les évolutions d'un type d'occupation des sols particulier, par exemple les *zones industrielles*. Dans ce cas, il est préférable de définir des 'zones industrielles' de façon indépendante des 'parcelles' sur lesquelles elles se trouvent. Cela permet de disposer d'une composante géométrique propre aux 'zones industrielles' et de pouvoir alors leur associer une composante sémantique pour porter les valeurs d'attributs spécifiques à chaque 'zone' (ex. : nombre d'entreprises, types d'activités, nombre d'emplois). La figure B.1.15 illustre ce point avec une autre type d'entités : les inscriptions sur un mur. Lorsque l'inscription sur un mur est considérée comme un *tag* vite effacé, l''objet-inscription' n'existera même pas. Dans le cas où elle est un *indicateur d'« incivilité »*, elle peut venir « incrémenter » l'attribut 'nombre de tags annuels' associé à un tronçon de rue ou à un bâtiment (pour une localisation précise). Dans le cas où elle est une « *opération politique »*, on peut la localiser sous forme de point pour faciliter le bilan annuel et le recensement des actions d'animations menées sur le territoire d'une commune (avec comme attribut associé par exemple le nombre d'articles évoquant l'événement). Enfin, *l'inscription-œuvre* d'art pourra être définie sous forme de ponctuel, de linéaire ou de surfacique suivant son emprise et la forme de son support. Si par ailleurs elle fait l'objet de mesures de protection, elle pourra servir à définir une 'zone' correspondant au *périmètre de servitudes*.

10. Le *géocodage* est l'opération qui permet d'associer une localisation à une information qui n'en possède pas en utilisant une relation existant entre cette information et un autre 'objet' déjà localisé. Par exemple, un individu peut être géocodé par son adresse, un accident par sa référence à un point routier localisé sur le réseau routier.
Le géocodage peut utiliser plusieurs méthodes techniques (ex. : la jonction de fichier, pour une localisation indirecte au sein des attributs d'un 'objet' existant), mais classiquement il évoque souvent la création d'une nouvelle primitive ponctuelle (pour une localisation directe) ; Dans ce dernier cas l'objet un fois géocodé, peut alors servir de support à d'autres informations (exemple, le logement géocodé au bâtiment qui permet de localiser ses habitants)

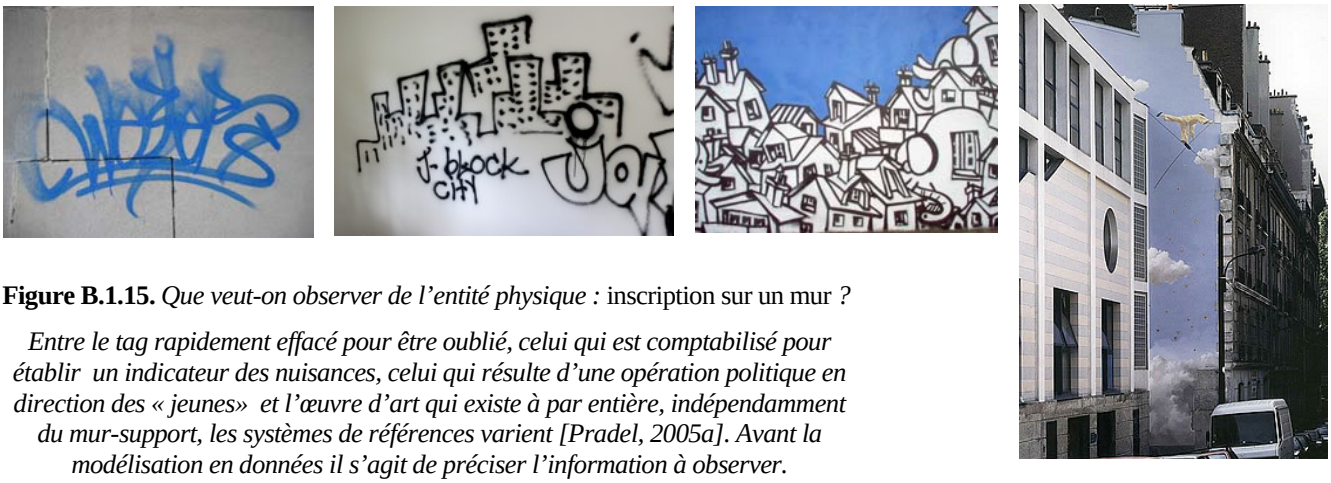


Figure B.1.15. *Que veut-on observer de l'entité physique :* inscription sur un mur *?*

Entre le tag rapidement effacé pour être oublié, celui qui est comptabilisé pour établir un indicateur des nuisances, celui qui résulte d'une opération politique en direction des « jeunes» et l'œuvre d'art qui existe à par entière, indépendamment du mur-support, les systèmes de références varient [Pradel, 2005a]. Avant la modélisation en données il s'agit de préciser l'information à observer.

Typologie des informations géographiques en fonction de la forme de la localisation

La localisation « fait » l'information géographique, mais d'un point de vue technique cette localisation peut prendre des formes différentes :

- *une localisation implicite* par l'intermédiaire d'un attribut descriptif faisant référence à une *commune*, un *pays*, une *adresse*, un *point kilométrique*, etc. (par exemple un *client* sera localisé par son adresse, un *ticket de caisse* par l'adresse du *client* à qui il appartient, un *accident* sera localisé par rapport à un *PK* ou *PR* point routier).
- *une localisation explicite indirecte* sous la forme d'attribut, où l'information sert à caractériser une autre information géographique (par exemple, le champ de blé peut être un attribut d'une information parcelle, une école un attribut d'un bâtiment)
- *une localisation explicite simple, ponctuelle*, obtenue généralement après l'opération de géocodage. Le géocodage permet en effet d'associer des coordonnées à des objets géographiques localisés implicitement en référence à un autre objet : lorsque la composante géométrique de l'objet de référence est un point, les coordonnées associées sont celles du point, lorsque l'objet de référence possède une composante géométrique polygonale, les coordonnées sont souvent celles de son centroïde, et pour les objets linéaires, les coordonnées sont obtenues par calculs barycentriques.
- *une localisation explicite plus élaborée* sous forme surfacique ou linaire correspondant à une emprise sur le sol, donnée par une composante géométrique de type polygone ou polyligne.

La forme de la localisation doit être fonction du rôle descriptif des informations géographiques [Didier, 1990]. Afin de préserver ce lien lors de l'implémentation en données géographiques et pour aider à la formalisation technique, nous proposons de distinguer trois types d'informations en fonction de leur localisation [Bordin, 2005b], [Bordin, 2002] et nous développons ensuite deux exemples pour montrer comment le choix d'un type contribue aux analyses thématiques. Ainsi, nous appelons respectivement :

Information localisable (localisation implicite): information qui n'a pas de coordonnées, mais qui possède un attribut la mettant en relation avec une donnée géographique dans un SIG (ex. : le 'client' et son adresse)

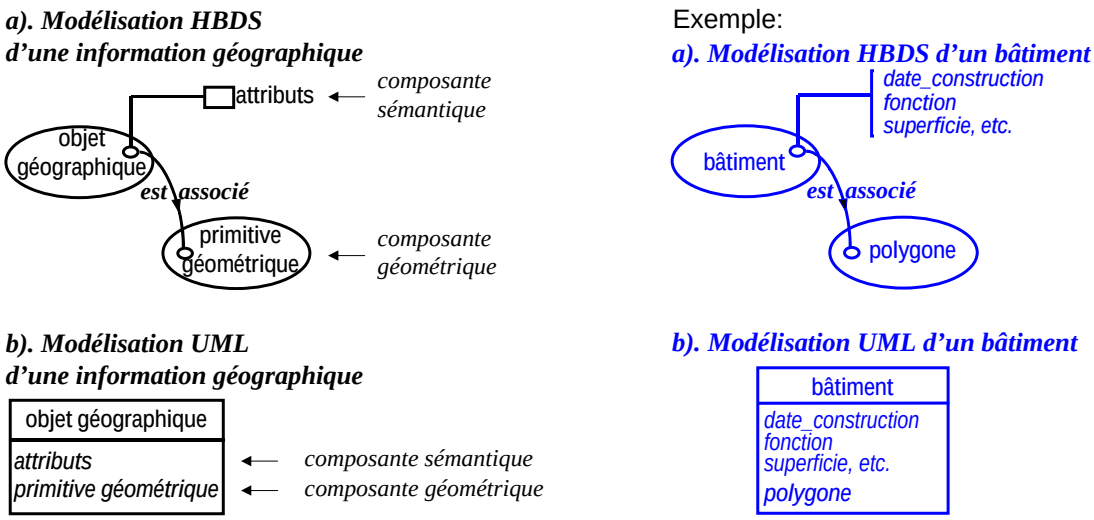
Information localisée (localisation explicite indirecte): information qui décrit sous forme d'attribut une autre information géographique dite (*information) support* (ex. : le type de culture d'un champ, le nombre d'élèves dans une école),

Information géographique par nature (localisation explicite directe)[11] : information qui possède une composante géométrique propre et qui peut ainsi devenir le support d'informations localisées (par enrichissement de la composante sémantique) ou servir à géocoder des informations localisables.

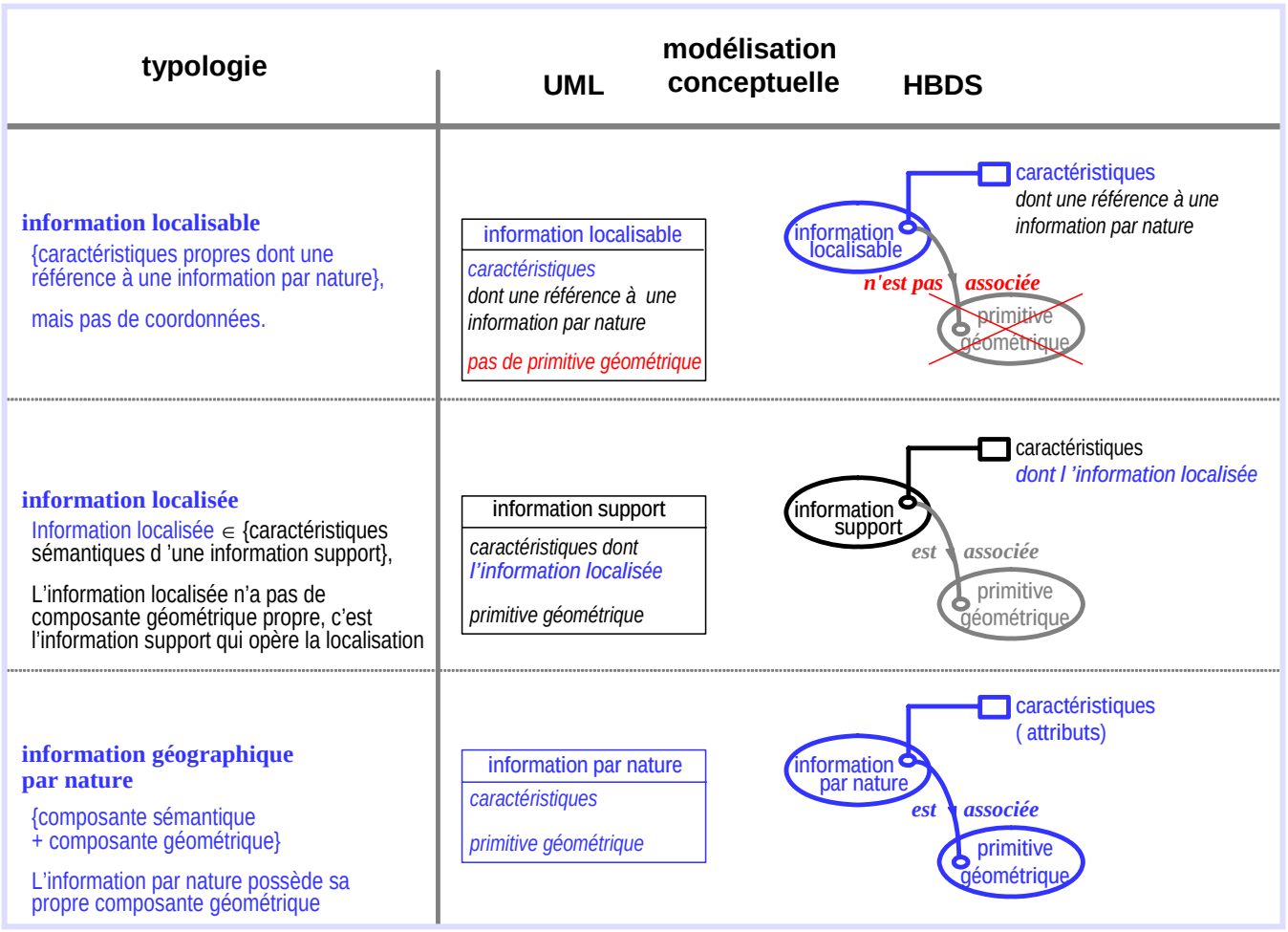
En informatique pour présenter le résultat d'une modélisation, on utilise un *modèle conceptuel*[12]. La figure B.1.16 présente les deux modèles exploités dans cette thèse : HBDS (développé par [Bouillé, 1977]) que nous utiliserons pour discriminer le rôle des composantes de l'information géographique et UML (modèle classique devenu un standard) que nous emploierons pour présenter des éléments techniques sur la méthode proposée. Dans la figure B.1.17, ils nous servent à montrer la différence de modélisation entre les trois types d'informations géographiques.

11. Nous avons regroupé dans la même catégorie toutes les informations dont la localisation est explicite (simple et élaborée), car la différence ne reflète que des distinctions techniques qui portent :
- soit sur les processus d'acquisition (géocodage ou mode d'acquisition plus complexe comme la digitalisation, la photogrammétrie),
- soit sur l'échelle de représentation (ex. : on peut digitaliser un bâtiment comme un point à une certaine échelle et comme un polygone à une autre),
et qui n'intéressent pas directement le thématicien, mais concerne le technicien dans la phase d'élaboration des données.

12. Modèle conceptuel : Méthode de représentation graphique (ex. : Merise, UML, HBDS) proposant des règles et des contraintes de représentation. Il se distingue du schéma conceptuel qui est la déclinaison de ces règles générales à un cadre applicatif particulier.



FigureB.1.16. *Deux modèles conceptuels différents pour représenter les informations géographiques HBDS (a) et UML (b)*



FigureB.1.17. *Modélisations conceptuelles comparées des différents types d'informations géographiques*

EXEMPLES :

- Les données de l'INSEE sur les îlots ont été conçues comme des informations localisables. En effet, les données sémantiques peuvent être livrées indépendamment de la composante géométrique. En revanche, lorsque l'on joint les deux composantes de l'information géographique 'îlot' à l'aide de l'identifiant fourni, l'îlot' devient une information par nature.

- Dans le MOS, les 'zones d'habitat individuel' sont des informations localisées (voir l'explication détaillée page suivante)

- Les informations sur les bâtiments dans la BD TOPO® ont été conçues comme des informations par nature et leur composante géométrique est décrite sous forme de polygone. Toutefois, on note ici que leur composante sémantique ne contient pas beaucoup d'attributs .

Dans les SIG, la modélisation sous forme d'information géographique par nature est la plus complète et la plus riche. Il est essentiel qu'elle soit retenue pour les objets principaux de l'observation. En effet :

– le fait de disposer d'une composante géométrique explicite

permet d'effectuer des analyses de type morphologique et/ou topologique, qui peuvent contribuer à une plus grande connaissance de l'*objet*,

– le fait de disposer d'une composante géométrique directe

- permet de localiser les informations de façon immédiate avec la requête élémentaire *où* ? (la visualisation des informations localisables implique quant à elle une étape de géocodage préalable qui peut être longue et laborieuse en fonction de la qualité des données exploitées).

- permet d'associer à l''objet' une composante sémantique propre que l'on peut ensuite enrichir « à l'infini », ce qui n'est pas le cas des informations géographiques localisées, comme nous allons le voir maintenant sur le cas des 'zones d'habitat individuel' dans le MOS. Ce qui offre la possibilité d'analyses statistiques plus nombreuses.

– le fait de disposer d'une composante géométrique et d'une composante sémantique permet de faciliter les observations simples (de type inventaire, recensement, localisation) dans les SIG (cf. Annexe 4).

B.1.2.3. Informations à observer, informations observables : rôle de la modélisation en 'objet technique' illustré sur deux exemples

EXEMPLE 1 : *Les 'zones d'habitat individuel' dans le MOS*

La modélisation du MOS se fonde sur une partition du territoire en 'parcelles', unités géographiques d'occupation des sols homogènes. Les 'parcelles' sont de fait des informations géographiques par nature (ce sont d'ailleurs les seules que contient le MOS). Par ailleurs, le MOS permet d'observer 83 types d'occupations des sols différents. Ces types sont implémentés sous forme de code (ex. : '41' pour l''habitat individuel')[13]. Ces codes constituent les valeurs possibles pour les cinq attributs que comporte la composante sémantique d'une 'parcelle' ; ces attributs correspondant chacun à l'état de la parcelle à une date donnée (fig. B.1.18 et B.1.19).

13. Le choix d'implémenter un code s'explique par des raisons techniques : il est plus rapide, plus simple et plus sûr de rentrer un code au lieu d'une chaîne de caractères
ex. : « 41 » au lieu de :
« zone_habitat_individuel »

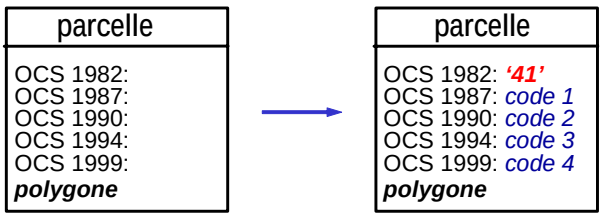


Figure B.1.18. *Modélisation UML de l'information géographique par nature 'parcelle' et de l'information géographique localisée 'zone d'habitat individuel en 1982' dans le MOS*

L'information 'parcelle' comporte une composante géométrique de type polygone et une composante sémantique constituée de 5 attributs 'OCS (occupation des sols)' à 5 dates différentes. L'information 'zone d'habitat individuel en 1982' est représentée sous la forme de la valeur '41' pour l'attribut 'OCS en 1982' de l'information 'parcelle' (les valeurs des autres attributs pouvant prendre n'importe quelle valeur parmi les 83 spécifiées dans le MOS)

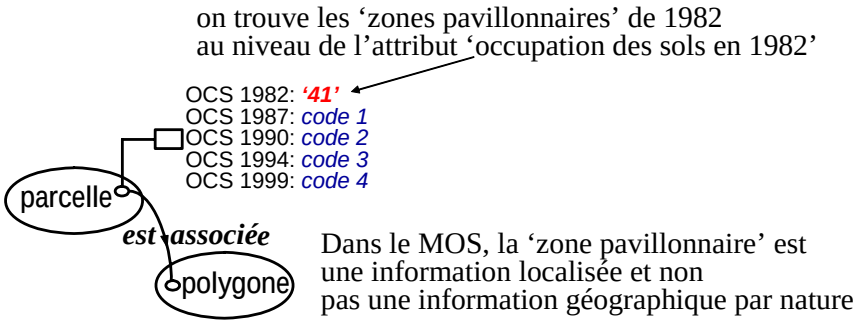

on trouve les 'zones pavillonnaires' de 1982
au niveau de l'attribut 'occupation des sols en 1982'
OCS 1982: '41'
OCS 1987: code 1
OCS 1990: code 2
OCS 1994: code 3
OCS 1999: code 4
parcelle
est associée
polygone
Dans le MOS, la 'zone pavillonnaire' est
une information localisée et non
pas une information géographique par nature

Figure. B.1.19. *Modélisation HBDS de l'information ' zone d'habitat individuel en 1982' dans le MOS*

Si l'on veut observer les *zones d'habitat individuel (ZHI) de 1982*, il faut faire une requête sur l'attribut des 'parcelles' correspondant à la date 1982 : *« sélectionner les parcelles telles que code d'occupation des sols en 1982= '41' »*. On obtient alors un sous-ensemble de 'parcelles' et non des ZHI. En effet, si on veut caractériser les ZHI par leur nombre de bâtiments, il n'existe pas d'autre solution que de créer un attribut 'nombre de bâtiments en 1982' et de l'associer à la 'parcelle'. En faisant ainsi, les parcelles de type 'grande infrastructure', 'berge' ou encore 'verger' posséderont également cet attribut (ce qui n'a pas de sens géographique). Pour pouvoir caractériser les ZHI, il faut qu'elles soient sous forme d'informations géographiques par nature. Dans le cas du MOS, il est simple de les établir comme telle : il suffit de sélectionner pour chaque date l'ensemble des 'parcelles' de type 'ZHI', de dupliquer leur géométrie et de définir les informations ZHI par nature à partir de ces composantes géométriques dupliquées. Une fois que les ZHI possèdent leur propre géométrie, il ne reste plus qu'à créer les attributs permettant de recevoir les informations descriptives que l'on veut associer aux ZHI (fig. B.1.20).

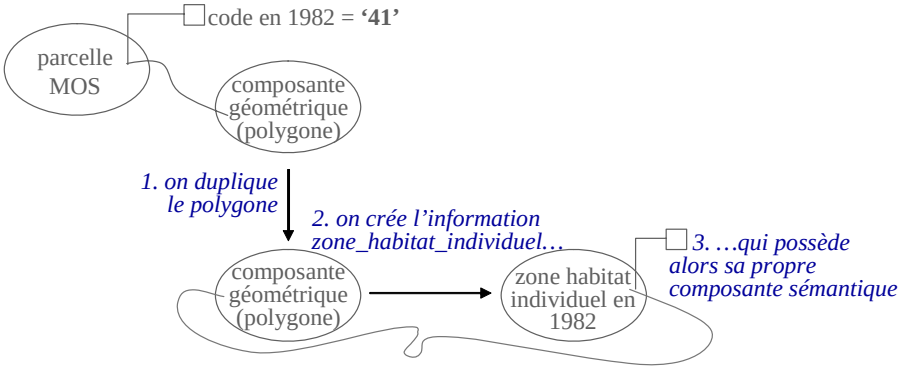


Figure B.1.20. *Création de l'information géographique par nature :'zone d'habitat individuel' à partir des données sur les parcelles*

Il est importe ici de noter que la sélection des 'ZHI' se faisant par année, la composante géométrique d'une 'parcelle' peut être dupliquée plusieurs fois (autant de fois qu'il existe d'années de référence où la 'parcelle' est 'ZHI'). Les parcelles de type pavillonnaire *depuis* 1982 engendreront ainsi cinq 'ZHI' différentes (une pour chacune des dates) qui n'auront *aucun* lien informatique entre elles. Contrairement aux parcelles dans le MOS, que l'utilisateur peut suivre facilement grâce à la structure spécifique de cette base, *le suivi automatique des 'ZHI' ainsi construites ne pourra se faire qu'à l'aide de développements spécifiques*[14].

14. Ces développements peuvent porter sur la mise en place d'outils d'appariement. Mais, il existe une alternative plus simple qui consiste à attribuer un identifiant à chaque parcelle (elles n'en possèdent pas), puis de créer un attribut 'identifiant de la parcelle d'origine' lors de la duplication de la composante géométrique. L'identifiant de la 'parcelle' associée à la 'date de référence' des ZHI permet alors de suivre la parcelle au travers des ZHI successives... sauf si à l'une des dates la parcelle change d'occupation des sols. Dans ce cas, le suivi de la parcelle est interrompu.

EXEMPLE 2 : *Les 'zones d'accidents routiers'*

La gestion des accidents est une application qui donne lieu à des recherches diverses : certaines travaillent à identifier les *sites dangereux* [Huguenin-Richard, 2000], d'autres cherchent à établir des liens entre les *accidents* et leur contexte [Millot, 2003]. Schématiquement, elle repose sur la volonté d'identifier les *sites dangereux* et d'en comprendre l'origine. Par exemple, on cherche à connaître les *tronçons dangereux* (avec des études de localisation relative des *accidents* par rapport aux *tronçons* faisant appel à la topologie et à la géométrie) et à étudier les causes des *accidents* (avec des études statistiques sur les horaires, le type de véhicule, le taux d'alcoolémie, etc.). Il ne s'agit donc pas seulement de les visualiser et de les localiser grâce à une représentation cartographique, mais bien d'effectuer des analyses à référence spatiale. Cependant partant du besoin de représentation visuelle, la modélisation du problème est souvent influencé par l'expérience cartographique. En particulier, le nombre de variables étant limité en cartographie (entre autre pour des raisons de lisibilité), il arrive que l'on choisisse d'agréger certaines informations, alors que ce choix ne s'impose pas dans les SIG. Ainsi, dans le cas des *accidents*, on peut décrire les 'tronçons de route' par le 'nombre d'accidents' survenus en un an sur la portion de route (cf. la carte fournie en 1992 par la DDE de Haute Garonne pour illustrer les résultats obtenus grâce au SIG fig. B.1.21). Cette description permet d'identifier rapidement les *tronçons* qui ont subi plusieurs accidents.

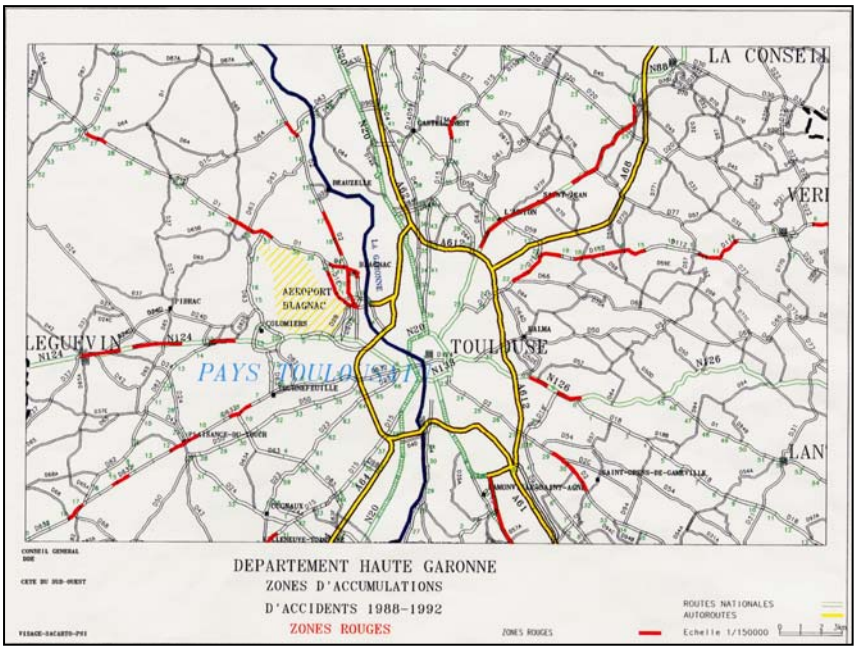


Figure B.1.21. *Carte des accidents représentés par tronçons*

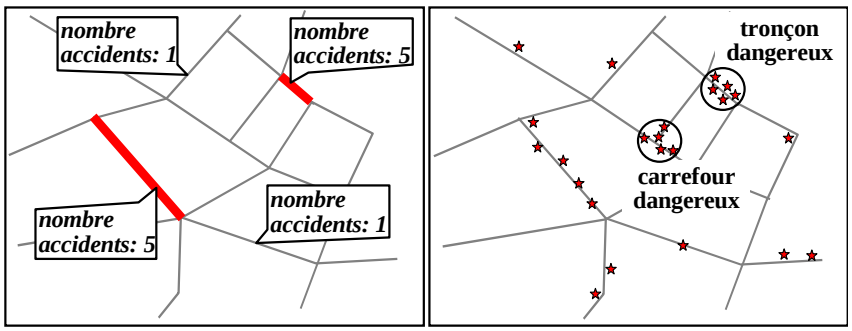


Figure B.1.22. *Biais dû à la représentation cartographique des accidents sous forme de nombre cumulé par tronçon*

Cependant la figure B.1.22 met en évidence certains biais de cette représentation, bien connus des cartographes [Thomas, 1996] : la modélisation des 'accidents' sous forme d'un nombre d'accidents cumulés associé à chaque 'tronçon' peut engendrer de mauvaises interprétations lors de l'analyse visuelle (pour un nombre égal d'accidents, un 'tronçon' de grande longueur paraît aussi dangereux qu'un 'tronçon' court alors qu'il présente une concentration moins grande d'accidents). Par ailleurs cette représentation, où l'information 'accident' est localisée sur les 'tronçons' ne permet pas d'identifier les 'carrefours dangereux'. Pour remédier au premier biais, on peut réaliser un changement de variable et considérer le nombre d'accidents par mètre linéaire. Pour remédier au second, la solution consistant à introduire les 'carrefours' semble séduisante, mais elle n'est pas entièrement satisfaisante. Tout d'abord, elle fonctionne moins bien que celle sur les 'tronçons'. En effet, pour prendre les 'carrefours' comme information support, il faut que ceux-ci possèdent une composante géométrique. Or, l'information 'carrefour' n'existe pas toujours sous forme d'information par nature de type polygone pour permettre d'agréger les 'accidents'. Il s'agit donc au préalable de construire ces informations. Toutefois, les limites des *carrefours* sont plus floues que celles d'un *tronçon de route*. Il peut en résulter des définitions de 'carrefours' variables suivant les utilisateurs (fig. B.1.23). Par ailleurs, les défauts d'une modélisation des 'accidents' sous forme d'information localisée (aux 'carrefours' et /ou aux 'tronçons') ne se limitent pas aux biais cartographiques présentés. Plus fondamentalement, cette modélisation restreint les analyses à référence spatiale que l'on peut mener.

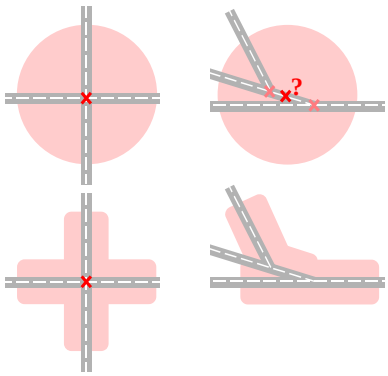


Figure B.1. 23. *Différentes possibilités pour définir un 'carrefour'*

En haut en considère un buffer autour du croisement, en bas autour des tronçons de route, à gauche on n'envisage qu'un croisement, à droite on envisage qu'il puisse y avoir plusieurs croisements.

L'observation des *zones dangereuses* ne se limite pas à leur recensement et à leur localisation : il s'agit aussi de rechercher leurs causes pour essayer d'intervenir. Dans la pratique, le fichier CONCERTO qui comporte les données sur les accidents normalisées à l'échelle européenne propose près de 300 valeurs d'attributs réparties en plusieurs catégories (ex. : conducteur {age, sexe, taux d'alcoolémie}, véhicule {type de véhicule, année de construction, immatriculation}, circonstances {temps, heure, jour}, etc.). Cependant, une modélisation des 'accidents' comme *information localisée* empêche toute étude de ces caractéristiques, puisqu'il n'y a pas de composante sémantique pour les intégrer. Une modélisation sous forme *d'information localisable* permet de travailler sur l'ensemble des attributs des 'accidents', en revanche elle ne les localise pas explicitement. Elle peut servir pour des études statistiques générales sur les *accidents*, mais pas pour dégager les spécificités de chaque *zone dangereuse*. En effet, les 'accidents' n'étant pas localisés, on ne peut pas identifier ceux qui se trouvent à proximité d'autres 'accidents', on ne peut donc pas définir de 'zones dangereuses', ni identifier parmi les 'accidents' ceux qui appartiennent à une même 'zone', et par conséquent on ne peut pas rechercher les caractéristiques communes aux 'accidents' d'une même 'zone dangereuse'.

Seule la modélisation des accidents sous forme *d'information par nature* permet à la fois de définir des 'zones dangereuses' et d'identifier les particularités des 'accidents' appartenant à chaque zone. (Par exemple, les accidents ont lieu essentiellement le samedi soir en fin de nuit ou le dimanche au petit matin, sur des jeunes adultes ou encore plutôt le vendredi soir à la tombée de la nuit). Ces informations de détails donnent alors des éléments pour rechercher les causes, par exemple en introduisant des informations complémentaires comme les discothèques au voisinage ou en effectuant une enquête sur le terrain pour essayer d'identifier une cause très particulière (ex. : une implantation d'arbres qui fait effet de stroboscope au coucher du soleil provoquant l'endormissement des conducteurs fatigués à la fin de la semaine) (fig. B.1.24) (cf. figure B.1.25).

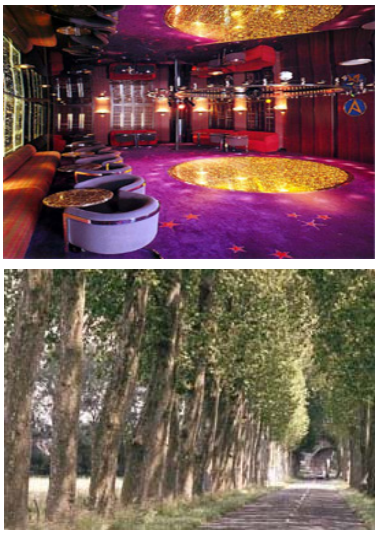
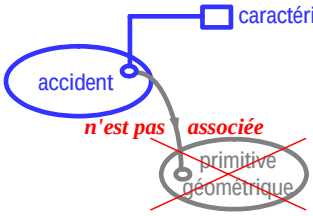
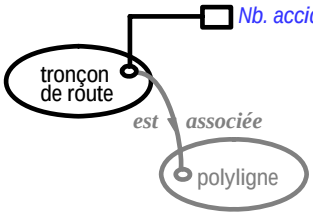


Figure B.1.24. *L'étude des zones d'accidents passe par leur identification sous forme de 'zone de concentration d'accidents', puis par la recherche de caractéristiques communes aux 'accidents' d'une même zone. L'ultime étape consistant à interpréter ces similarités pour tenter de trouver l'origine commune des accidents .*



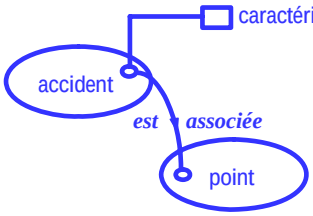
Les accidents sont décrits par l'ensemble de leurs caractéristiques sémantiques.
On peut ainsi effectuer des statistiques sur les accidents, en général.
En revanche, comme ils ne sont pas localisés (pas de composante géométrique associée),
on ne peut pas voir ceux, qui par proximité, constituent un groupe une 'zone dangereuse'.

⇒ *Ce modèle permet de faire des études statistiques générales sur les accidents*



L'accumulation des accidents sur un 'tronçon' permet de définir des 'zones dangereuses'
sous forme de 'tronçon de route'. Réciproquement, on peut identifier les 'zones dangereuses',
si elles correspondent effectivement à un 'tronçon de route'.
Pour identifier celles correspondant à un 'carrefour', il faut définir un objet 'carrefour'
sous forme d'information par nature avec un attribut 'nombre d'accidents'.
En revanche, aucune information pour caractériser les accidents ne peut être saisit avec ce modèle.

⇒ *Ce modèle permet de localiser le danger (relativement à des 'objets' préétablis),*
mais ne permet pas de rechercher simplement l'origine du danger.



Les accidents existant sous forme d'information par nature, ils sont localisés.
On peut alors facilement identifier les zones d'accumulation et les définir comme des 'zones dangereuses'.
Par requête sur les caractéristiques des accidents de chaque zone, on peut tenter d'identifier
les caractéristiques communes. Ce qui donne des éléments pour rechercher l'origine du danger
(qui peut n'avoir aucune relation de causalité directe avec le 'tronçon de route' ou le 'carrefour' le plus voisin).

⇒ *Ce modèle permet de définir des zones dangereuses, sans devoir se référer à des objets existants.*
Il permet également d'étudier les caractéristiques des accidents participant spécifiquement
à une même zone pour tenter de mieux comprendre la cause de cette 'zone'.

Figure B.1.25. *Choix de modélisation des 'accidents' en information géographique et conséquences pour les analyses à références spatiales ultérieures*

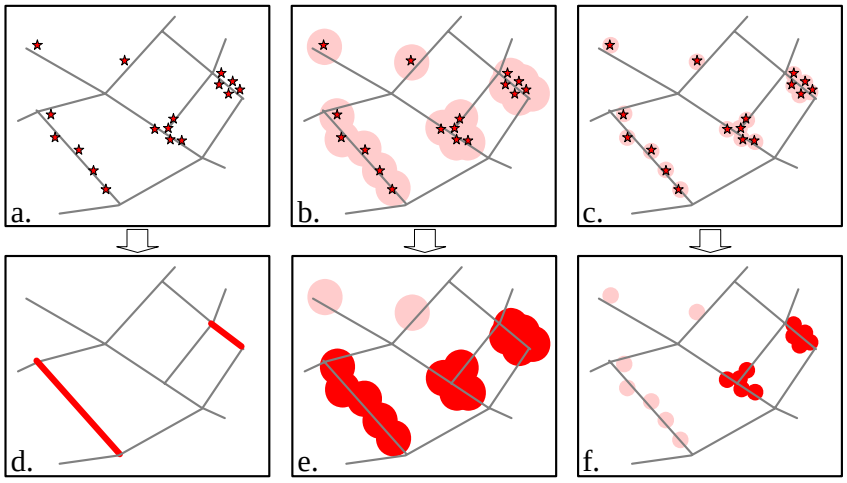


Figure B.1.26. *Définition de l'objet d'observation principale: les 'zones dangereuses'*

Cette application qui porte sur l'observation des 'zone d'accidents routiers' met en avant un certain nombre de points :

– elle montre que l'étape de formalisation technique peut à elle seule introduire un décalage entre les informations que l'on veut observer et les informations observables dans le SIG. Ainsi, pour étudier les 'zones dangereuses', il existe plusieurs façon de les définir. Par exemple à partir d'un seuil limite d'un attribut 'nombre d'accidents' d'un objet géographique support, par exemple les 'tronçons de route' (fig. B.1.26a et d). Cependant, l'expérience des 'carrefours' montre que cela peut ne pas suffire.

En fait, cette expérience conduit à soulever une question plus fondamentale : en choisissant un 'objet' géographique pour servir de support à une information, on effectue un lien implicite en considérant que le phénomène caractérise effectivement cet *objet* ; Or dans le cas des 'accidents' localisés sur les 'tronçons', si l'exemple sur l'alignement d'arbres monte que ce type de lien peut exister, l'exemple sur la sortie de discothèque montre que ce n'est pas toujours le cas (ce ne sont pas les spécificités du *tronçon* qui sont a l'origine des accidents mais des éléments de contexte, les circonstances (heure, conditions de conduite, distance par rapport à la discothèque et probablement au domicile, etc.). Ainsi, peut-il être plus judicieux de définir les 'zone dangereuses' à partir de l'agrégation des accidents' qui servent à les révéler et à les caractériser. La figure B.1.26 (b-e et c-f) montre que cela soulève par ailleurs des questions de spécifications techniques (ex. : définition de la taille, de la forme du buffer, de l''objet' de référence). Ces éléments de définitions qui requiert la collaboration entre thématicien et technicien pour décider de la « meilleure » solution font partie de la mise en œuvre des applications spécifiques et ne peuvent en revanche pas être déterminés dans une méthodologie générique.

– cette application montre les enjeux d'une formalisation sous forme d'*information géographique par nature* :

- cette modélisation permet d'observer directement, à l'aide de requêtes élémentaires, des aspects à la fois de localisation (contrairement aux *informations localisables*) *et* de caractérisation (contrairement aux *informations localisées*). Cette modélisation doit donc être privilégiée pour les objets d'observations principaux. Ainsi, lorsque l'on identifie les 'zones dangereuses' aux 'tronçons de route' qui existent déjà sous forme d'information par nature, plutôt que d'observer l'ensemble des 'accidents', on préfère souvent créer un attribut 'nombre d'accidents' qui transforme l''accident' en information localisée, mais permet d'identifier immédiatement les 'zones dangereuses' ('tronçons dangereux') par simple requêtes sur la valeur de cet attribut (ex. : valeur > 5).

- cette modélisation permet d'effectuer des analyses plus élaborées que les requêtes élémentaires, sur des aspects de localisations et sur les caractéristiques sémantiques[15]. Il s'agit donc de la privilégier pour les *objets* sur lesquels on veut effectuer des analyses à référence spatiale. Ainsi, le fait de modéliser les 'accidents' sous forme d'informations géographiques par nature permet d'étudier les relations de voisinage pour définir par agrégation des 'zones dangereuses' plus pertinentes. Il autorise par ailleurs l'utilisation d'outils statistiques pour rechercher les corrélations existant entre les différentes caractéristiques des *accidents* d'un point de vue général (ex. : taux d'alcoolémie et sexe, ou age, type de voiture et age, etc.) ou chercher les caractéristiques communes aux 'accidents' décrivant une 'zone particulière' (ex. : tranche horaire),

15. En revanche, nous regrettons que l'application n'ait pas permis de mettre également en valeur ici les avantages exclusifs de cette représentation pour étudier les aspects morphologiques et topologiques.

– enfin, cette application montre de façon fortuite, qu'il peut être intéressant d'introduire deux niveaux d'observation :

- un niveau principal, pour identifier rapidement et simplement (par requêtes élémentaires) les objets à observer principaux (et leurs caractéristiques), ici les 'zones dangereuses',

- et un autre niveau, plus fin, sur lequel on réalise des analyses à références spatiales pour élaborer des informations sur le niveau principal, en vue d'une meilleure connaissance voire d'une meilleure compréhension du phénomène.

B.1.3. Vers l'observation de phénomènes géographiques avec un SIG

Ce chapitre montre l'importance de l'étape de formalisation technique qui fait le lien entre la formalisation thématique et l'implémentation. Il distingue ensuite trois types de modélisation en information géographique et désigne parmi eux, la modélisation sous forme d'« information géographique par nature[16] » comme le type à privilégier :

- pour une observation simple et facile à enrichir,
- pour des analyses à référence spatiale plus nombreuses et plus variées.

L'exemple développé sur les *zones de circulation dangereuses* et l'expérience d'observations visuelles sur des *zones pavillonnaire* (cf. p. 89) suggèrent la mise en place d'une observation à deux niveaux. Cependant dans le chapitre suivant, nous verrons que pour pouvoir la mettre en œuvre, il nous faut dépasser une difficulté technique supplémentaire dans les SIG : la gestion simultanée de représentations différentes.

16. Par la suite, il pourra nous arriver d'abréger l'appellation « information géographique par nature » en l'expression «information par nature ».

B.2. Méthode d'observation multi-niveaux dans un SIG et méthode d'agrégation par emprise

B.2.1. Observation statique de phénomènes géographiques avec un SIG

Dans le domaine des SIG, il existe peu d'éléments méthodologiques pour mettre en œuvre des applications de façon générique. Le chapitre précédent explique pourquoi : la plupart des étapes participant au processus de modélisation des entités en données géographiques demandent en effet un travail à façon spécifique à chaque application. Dans cette thèse, nous voulons développer une méthode générique pour suivre des phénomènes géographiques à l'aide d'un SIG. Comme nous le précisons dans le chapitre A.2, lorsque nous envisageons de « *suivre un phénomène géographique* » nous ne projetons pas d'observer le processus en jeu, mais d'étudier ses effets ; nous proposons d'identifier et d'observer des *objets* associés au phénomène étudié et d'analyser leurs évolutions. Il s'agit d'appréhender le phénomène à partir d'informations sur la nature et la forme de ces *objets*. Par exemple, dans un phénomène d'épidémie les caractéristiques qui nous intéressent ne portent pas sur le processus de propagation en tant que tel (ex. : type de transmission, modalités de transmission, etc.) mais sur les territoires concernés par l'épidémie, sur le nombre de malades touchés, etc.

Le chapitre A.3 montre que, indépendamment de la difficulté intrinsèque à la définition d'une méthodologie générale, le type d'application retenu nous confronte à un problème technique complexe et encore délicat à résoudre dans les SIG : la gestion de la dimension temporelle. Des solutions existent pour le surmonter mais la plupart sont informatiques et elles peuvent être relativement compliquées à mettre en œuvre pour un non développeur.

Dans ce chapitre, nous nous concentrons sur la définition d'une méthode d'observation statique qui soit adaptée à l'observation thématique des phénomènes. L'aspect cinématique de l'observation fera l'objet du chapitre B.3.

B.2.1.1. Observation d'objets géographiques

Pour envisager une méthode d'observation statique pour les SIG, nous partons de deux idées pratiques :

– la nécessité de données à observer

– l'utilité d'être au plus près de la formalisation thématique

Le besoin de données

Sans données, il ne peut y avoir d'application dans un SIG. Or l'acquisition de données nouvelles reste un processus coûteux (en temps, en moyens et en compétences) ; ainsi, l'utilisateur n'a pas toujours la possibilité de produire celles dont il aurait besoin. Il doit se contenter de celles qui sont accessibles et dans les faits, les informations disponibles sont souvent insuffisantes : la composante sémantique n'est pas assez riche pour un suivi détaillé ou les attributs qu'elle contient ne conviennent pas, quelquefois la modélisation n'est pas adaptée, parfois même il n'y a aucune information directe sur l'*objet* à observer. Les utilisateurs sont ainsi régulièrement confrontés au manque de données sur les phénomènes qu'ils souhaitent étudier (ce qui est le cas des trois exemples d'applications que nous avons choisis) et l'observation d'un phénomène dans un SIG reste tributaire des informations géographiques disponibles sur le phénomène.

⇒ Nous proposons ici de développer une **méthode pour produire des informations dans un SIG,** pour permettre à l'utilisateur de **disposer de données** plus nombreuses et/ou mieux adaptées pour observer les phénomènes qu'il veut étudier.

La formalisation thématique d'une observation

Le chapitre B.1 a montré que pour aider les utilisateurs à observer les *objets géographiques* (et les caractéristiques) qui les intéressent avec un SIG, il importait que la formalisation technique en 'objet géographique' soit proche de la formalisation thématique. Or chaque phénomène géographique s'appuie sur une modélisation particulière. Il nous faudrait pourtant partir d'une formalisation thématique commune pour conserver à la méthode son caractère général. Les géographes qui travaillent sur l'analyse des phénomènes géographiques ont mis au point et proposent des méthodes d'analyses génériques.

⇒ Nous proposons donc de partir de la formalisation thématique de l'objet d'observation principal telle qu'elle est proposée par les géographes.
Dans un premier temps, il s'agit d'abord d'être proche d'une formalisation thématique générique des *objets géographiques,* pour spécifier les 'objets informatiques' et tenter de réduire le décalage à l'*entre-deux*. Dans un second temps, nous verrons qu'en retour, cela nous aidera à élaborer une méthode de création d'informations.

B.2.1.2. Définition de l'objet d'observation principal et niveau d'observation

Pour observer un phénomène les géographes définissent un objet principal d'observation, mais ils se demandent également : *«Quel est le niveau d'observation adéquat par rapport à la question à résoudre ? Quels sont les passages possibles, d'un point de vue sémantique et méthodologique, entre des niveaux différents ? »*. Ainsi, *« la première étape de nombreux travaux de recherches consiste à s'interroger sur le meilleur niveau d'observation à adopter relativement à l'objectif de l'étude entreprise »*

[Sanders, 2001, p.26] (fig. B.2.1). Par ailleurs, nous notons que les géographes envisagent également la possibilité de *« passages… entre des niveaux différents »* : *« Dans un certain nombre de cas, il ne s'agit pas d'opérer un choix mais, au contraire, d'utiliser simultanément plusieurs niveaux dans une perspective heuristique »* [Sanders, 2001, p. 27]. Les techniciens quant à eux sont plus réticents à envisager cette hypothèse, l'idée évoquant le « spectre » de la gestion de représentations multiples. Travailler dans un SIG *simultanément* avec des représentations différentes est en effet techniquement complexe, et ce quelque soit la différence, qu'il s'agisse :

- d'une différence de « niveaux d'observation »,
- d'une différence de points de vue thématique,
- ou même d'une « simple » différence de producteur des données géographiques, les points de vue et les « niveaux d'observation » étant égaux par ailleurs (cf. figure.B.2.2).

Pour pouvoir observer un phénomène géographique dans un SIG, il ne suffit pas de formaliser l'*objet principal* d'observation sous forme d'information géographique par nature ; il faut également intégrer dans cette formalisation la notion de « niveau d'observation ».

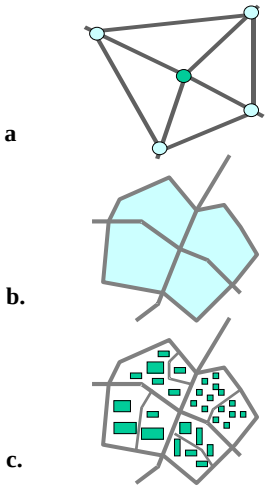


Figure B.2.1. *Trois niveaux de représentations différents pour un même objet*

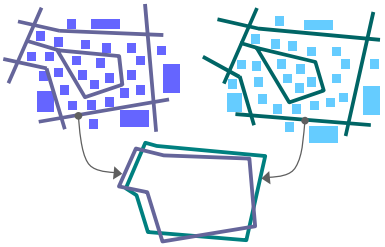


Figure B.2.2. *La numérisation d'un même 'objet' par des producteurs différents*

La numérisation d'un même 'objet' par des producteurs différents entraîne des représentations différentes ; les composantes géométriques ne se correspondent pas exactement.

Le niveau d'observation et les niveaux de détails

Le choix du niveau d'observation ne peut pas être générique. Sa formalisation dans un SIG peut en revanche utiliser des moyens généraux. Pour aller dans ce sens, nous proposons de rapprocher la notion de « niveau d'observation » qui n'existe pas en tant que telle dans les SIG, de la notion d'« échelle d'observation » pour laquelle il existe des éléments de formalisation. [Ruas et Mustière, 2005], [Ruas, 2002] proposent en effet des critères pour caractériser la façon dont le terrain observé est modélisé sous forme de données. Ces critères regroupés sous le terme de *niveaux de détails* (sémantiques et géométriques) peuvent servir pour juger de l'adaptation d'une base pour une observation à une échelle donnée. Ils peuvent également être utilisés comme une grille pour spécifier le contenu d'une base[1]. Par exemple, la résolution sémantique spécifie l'unité de référence utilisée pour les valeurs d'un attribut (ex. : la valeur de l'attribut 'nombre d'accidents de l'objet 'zone dangereuse' suivant la granularité sémantique {unité, dizaine, centaine}). Toutefois la notion de *niveau de détails* ne suffit pas pour formaliser techniquement le *niveau d'observation* géographique. D'abord parce que les *niveaux de détails* caractérisent des données ; ils incluent des éléments propres à l'étape d'implémentation[2], ensuite parce que la notion de *niveau d'observation* nous semble comporter un élément d'information supplémentaire sur l'observation par rapport à l'*échelle d'observation* : la *granularité thématique*.

La granularité thématique comme complément aux niveaux de détails

La granularité thématique évoquée dans le chapitre A.2, p. 49 stipule l'*objet* à partir duquel les analyses s'effectuent principalement, le plus petit *objet* insécable étudié. L'*objet* analysé principalement peut en effet ne pas être l'*objet* d'observation principal, mais un *objet* plus petit. L'exemple des 'zones dangereuses' (p. 108) montre l'intérêt qu'il y a à introduire un niveau d'analyse plus fin. Ainsi, [Batty, 2001] analyse les *individus* pour étudier la *ville*, [Boffet, 2001] s'appuie sur les 'bâtiments' pour définir la 'ville' (cf. figure B.2.3).

1. En particulier, ils aident à formaliser les questions sur ce qui doit être sélectionné (au sens de ce que l'on retient dans la base de données) et sur la forme que doit prendre cette sélection.

2. L'Annexe 1 présente les trois critères participant aux niveaux de détails géométriques d'une base de données.

La granularité thématique dépend de l'observateur et du phénomène et ne peut donc pas être générique non plus. Des principes généraux peuvent en revanche servir à l'implémenter dans les SIG. En particulier :

après avoir pris pour *principe de représenter sous forme d'information géographique par nature l'objet principal d'observation* pour bénéficier d'une commodité d'accès à l'information,

on peut poser le *principe de formaliser sous forme d'information par nature les objets désignés par la granularité thématique*, pour profiter du potentiel d'analyse offert par les fonctionnalités des SIG (cf. p . 111),

De façon théorique, ces deux principes assurent à l'observateur de disposer de deux niveaux de détails pour effectuer des analyses dans le SIG : celui de *l'objet* d'observation principal et celui des *objets* qui participent principalement aux analyses. Toutefois, il ne suffit pas d'effectuer des analyses à deux niveaux de détails pour pouvoir observer deux représentations d'un *objet* simultanément,

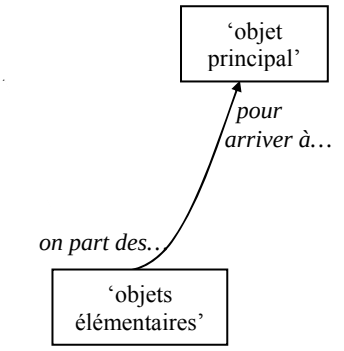
- l'une comme entité géographique sous forme d'information par nature
- et l'autre comme agrégat *d'objets* plus petits.

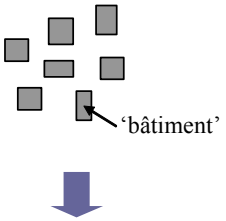
En particulier pour déduire de l'une (de l''agrégation') des caractéristiques pour l'autre (ex. : pour associer à la 'zone d'accident' la plage horaire d'occurrence maximale), il faut encore qu'il existe un lien entre les deux représentations (par exemple la représentation en information par nature des 'accidents' permet des analyses statistiques générales sur *les* accidents, mais une relation entre une 'zone dangereuse' et les 'accidents' est nécessaire pour préciser les caractéristiques spécifiques aux 'accidents' d'une 'zone'). Autrement dit, **dans un SIG il ne suffit pas de faire des analyses à des niveaux de détails différents pour pouvoir faire une analyse multi-niveaux** ; il faut également être capable de gérer conjointement les deux représentations dans le SIG, ce qui confronte le technicien au problème de gestion de représentations multiples. Dans l'exemple sur les accidents, il y échappait en représentant les 'accidents' comme un attribut des 'tronçons'. En effet, la modélisation sous forme d'information localisée établit un lien entre l'agrégat et l'objet support directe, mais ce lien existe alors au détriment de la capacité d'analyse.

Remarquons ici qu'une autre situation simplifie le problème des représentations multiples pour le technicien : quand il n'existe qu'*une seule* représentation ! C'est le cas lorsque l''objet' principal d'observation ne possède pas de représentation et que l'on décide alors de l'identifier à l'agrégation des 'objets' élémentaires (par exemple, dans la figure B.2.3, la 'ville' n'existant pas, on a choisi de l'identifier à l''emprise du bâti'). Le lien entre les deux représentations est alors déductif (l'un se déduit de l'autre) (cf. figure B.2.4). Le multi-échelles généralise ce lien (cf. B.2.2.1.1).

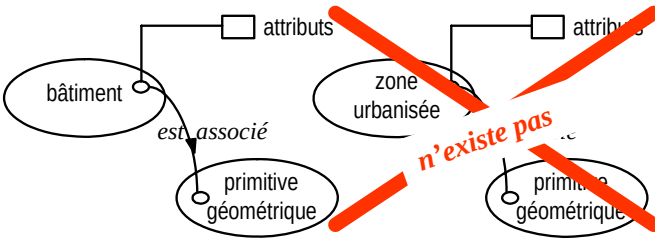
Figure B.2.3. *Observation d'une 'ville' à partir des 'objets' 'bâtiments'*

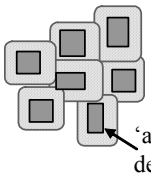
Pour observer l''objet' 'ville', [Boffet, 2001, p.102] exploite les 'objets' 'bâtiments' qu'elle « dilate » à l'aide de buffer de 25 m (a), puis elle prend l'enveloppe englobante des buffers agrégés (b) pour obtenir une représentation de la 'ville'.
Dans ce cas la représentation de l''objet principal' d'observation est déduite des 'objets élémentaires'.



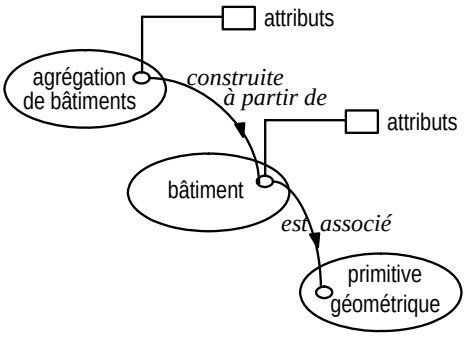


L'utilisateur dispose d'une représentation sous forme d'information par nature pour les 'bâtiments', mais pas pour les 'zones urbanisées'

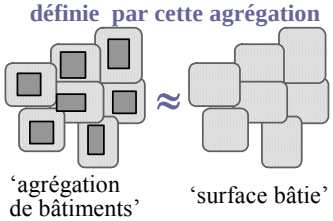




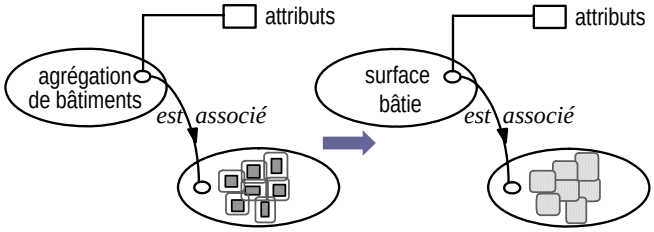
Dans le SIG, on peut construire à partir de fonctionnalités géométriques et topologiques (ex.: buffers) des 'agrégations de bâtiments'.



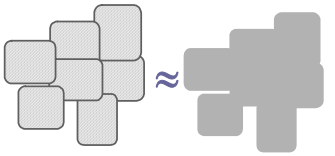
identification de l''agrégation de bâtiments' à la 'surface' définie par cette agrégation



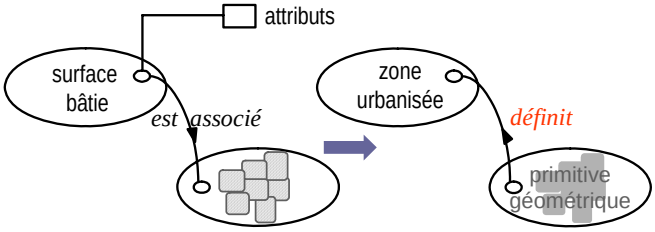
Définition de la 'surface bâtie' par identification des 'agrégations de bâtiments' à la surface qui leur est associée.



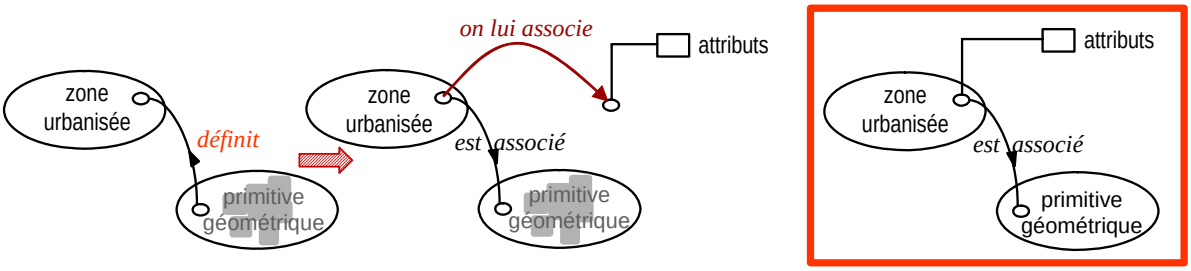
identification des emprises des *surfaces bâties* et des *zones urbanisées*



Définition des 'zones urbanisées' par identification des *zones urbanisées* aux *zones bâties*. Cela permet alors d'attribuer la composante géométrique de la 'surface bâtie' à la 'zone urbanisée'



La composante géométrique une fois créée, on peut lui associer une composante sémantique. On dispose alors d'une représentation sous forme d'information géographique par nature pour des objets 'zones urbanisées'.



L'observation conjointe des deux représentations (par nature et comme agrégation) est simple, puisqu'il y a identité de la primitive géométrique.

En particulier en s'appuyant sur cette identité, on peut transférer les caractéristiques calculées sur l'agrégation (ex.: surface moyenne du bâti) à la 'zone urbanisée' pour les observer comme l'une des caractéristiques de cette 'zone'.

Figure B.2.4. *Construction de la 'zone urbanisée' à partir d'une agrégation de 'zones bâties'.*

Quand l''objet' principal n'a pas de représentation sous forme d'information géographique par nature, on peut alors tenter de lui en définir une sous la forme d'une agrégation d''objets' d'un niveau plus fin, ce qui facilite l' observation multi-niveaux

B.2.2. Vers une observation multi-niveaux avec des SIG

Supposons résolu le problème de la gestion des représentations multiples dans les SIG. On pourrait alors effectuer des analyses multi-niveaux à partir de représentations à plusieurs niveaux de détails dans un SIG. Les caractéristiques calculées sur des 'agrégations' pourraient être utilisées pour décrire des 'objets' à un niveau plus général, pour une observation plus détaillées de ces 'objets' (par exemple, on pourrait décrire une 'zone pavillonnaire' à l'aide du nombre de 'bâtiments', de la surface moyenne de ces 'bâtiments', de leur forme, de leur date de création). On pourrait également introduire plusieurs types d''objets élémentaires' pour bénéficier d'une description plus variée. (Par exemple on peut choisir de caractériser un 'zone pavillonnaire à partir des 'bâtiments', mais aussi à partir des 'tronçons de routes', des 'parcelles', des 'habitants',… autant d'agrégations qui peuvent servir pour produire des descripteurs supplémentaires).

Ainsi, pour proposer une méthode générique d'observation statique de phénomènes géographiques avec un SIG, qui puisse produire des informations à observer, et qui soit proche de l'observation des thématiciens, il « ne reste plus qu'à » trouver une solution simple à implémenter pour les utilisateurs, qui leur permette de gérer conjointement deux représentations du même *objet* dans un SIG.

B.2.2.1. La gestion de multiples représentations dans les SIG

De leur origine cartographique, les SIG ont hérité une modélisation discrète des représentations en fonction des échelles ; les représentations correspondant à des échelles d'observation différentes existent indépendamment les unes des autres (dans un modèle équivalent pour l'espace, au *snapshots* pour le temps). Ainsi, si les outils de zoom avant/ arrière donnent l'impression d'un passage continu d'une représentation à une échelle à une représentation à une échelle proche, dans les faits, il s'agit toujours de la même base de données et un changement d'échelle plus marqué implique alors une rupture sous forme d'un changement de base de données. Il existe trois façons de concevoir les représentations multiples. La simplicité des solutions est fonction de leur performance :

– de la plus restrictive, où on ne s'intéresse qu'à l'*affichage* des diverses représentations, qui exploite des fonctionnalités standards des logiciels-SIG,

– à la plus élaborée : la *multi-représentations*, qui tente de faire cohabiter des bases de données variées pour permettre des études conjointes de représentations différentes sur un même *objet,* et qui nécessite des outils d'appariement,

– en passant par une approche intermédiaire, le *multi-échelles*, où l'on étudie conjointement des représentations du territoire à des niveaux d'observation différents, les représentations étant par ailleurs toutes déduites d'une même base de données, et qui implique des travaux de généralisation.

B.2.2.1.1. De la carte aux représentations multiples

Affichage multiple

La solution la plus simple pour gérer des représentations multiples répond uniquement à un besoin de type cartographique et *ne concerne que l'affichage*. Cette solution propose une simple *superposition des représentations*. Les représentations sont équivalentes aux calques en dessin. Dans un SIG, il existe pour cela des fonctions basiques : les bases de données ne font que cohabiter et c'est le logiciel de SIG qui gère leur affichage en fonction de la *gamme d'échelles d'utilisation*[3] qui leur a été préalablement associée. Suivant leurs fonctionnalités, les logiciels-SIG permettent de spécifier les échelles auxquelles certains groupes d'ʻobjetsʼ seront visibles ou invisibles (la « visibilité » généralement appelée *affichage* est ainsi spécifiée par *couche* dans MapInfo (fig. B.2.5a), par *type* ou *sous-type* dans Géoconcept). Certains logiciels proposent des options d'affichage plus sophistiquées qui agissent sur les classes d'ʻobjetsʼ et pas seulement sur un groupe plus large propre à la modélisation du logiciel. Ils permettent de définir pour une classe d'ʻobjetʼ quand les ʻobjetsʼ sont visibles et avec quelle symbolisation cartographique (voir les ʻBâtiments_Quelconqueʼ figure B.2.5b). La définition des classes d'ʻobjetsʼ est cependant soumise à la modélisation préalable dans le logiciel et dépend donc de l'étape d'implémentation.

3. La *gamme d'échelle d'utilisation* d'une base de données est le terme que nous avons proposé pour préciser les échelles « appropriées » à l'utilisation de la base. Cette notion représente pour les bases de données un équivalent de l'échelle pour une carte. Cependant, sa définition, empirique ne peut être aussi rigoureuse que celle de l'échelle.

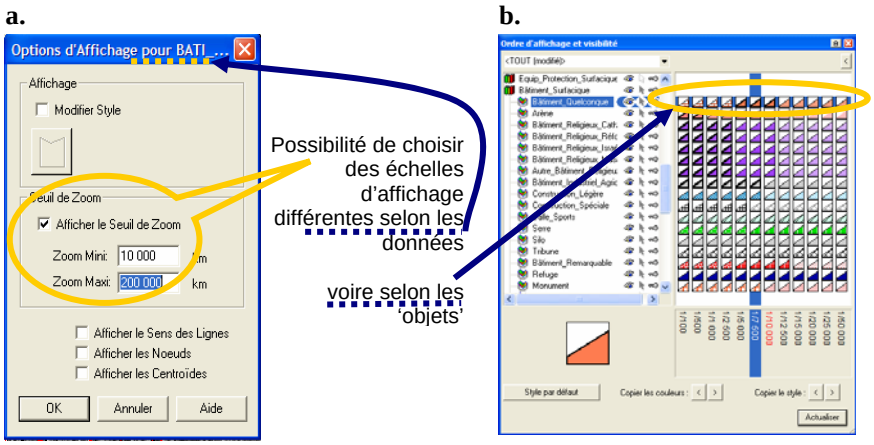


Figure B.2.5. *Fonctionnalités gérer l'affichage de représentations multiples*

Cette gestion des représentations multiples sous forme de bases de données autonomes qui cohabitent fait appel aux fonctionnalités d'*affichage* différencié des SIG. Elle peut convenir pour *gérer* des informations, lorsque les différentes représentations correspondent à des *thèmes* indépendants, chaque *objet* ayant généralement une seule représentation en ʻobjetʼ (cf. Annexe 2). Par exemple on représente les ʻbâtimentsʼ d'une part, les ʻroutesʼ d'autre part, le ʻréseau hydrographiqueʼ, etc. En revanche, elle n'est pas adaptée lorsque les diverses représentations correspondent à des niveaux de détails différents sur un même thème (autrement dit lorsqu'elles appartiennent à un même registre thématique). En effet dans ce mode de gestion, il n'existe aucun lien entre les différentes représentations d'un même *objet*, sauf celui que peut effectuer *visuellement* l'observateur, et qui donc ne peut pas être exploité de façon automatique.

Echelles multiples, déduction de représentations multiples

Dans une autre approche cartographique, la généralisation[4] propose de faire cohabiter des représentations différentes, mais en établissant un lien entre elles. Il ne s'agit pas ici de gérer l'affichage de représentations indépendantes. Il s'agit de *déduire d'une même base de données d'autres* base de données avec des niveaux de détails moins importants (c'est à dire comportant moins d'informations : moins de *thèmes* et/ou moins d''objets' dans les thèmes sélectionnés et/ou des 'objets' décrits moins précisément…). On parle de multi-échelles.

4. «*La généralisation cartographique est l'opération qui, par sélection schématisation et harmonisation, reconstitue sur une carte la réalité de la surface représentée dans ses traits essentiels en fonction du but de la carte, de son thème, de son échelle et des particularités de la région cartographiées*». [Cuenin, 1972, p.168].

Avec l'avènement de l'informatique, cela consiste souvent à déduire d'une représentation cartographique numérique une autre représentation à une échelle plus petite. Cette opération est complexe à automatiser car elle nécessite de sélectionner les informations retenues et donc de déterminer celles qui sont les plus pertinentes, mais aussi les plus représentatives, tant en terme thématique que de géométrie. Elle est l'objet de nombreuses recherches au Laboratoire COGIT.

Multi-représentations

Dans une approche plus analytique, d'autres travaux font coexister des représentations variées du territoire. Ils cherchent à mettre en relation les diverses représentations d'une « même » *entité* conçues de manière indépendante pour des applications différentes (ex. : applications stratégiques, tactiques ou opérationnelles s'appuyant sur des représentations plus ou moins synthétiques du territoire). On parle alors de représentations multiples ou de multi-représentations {Mustière, 2005} (figure B.2.6).

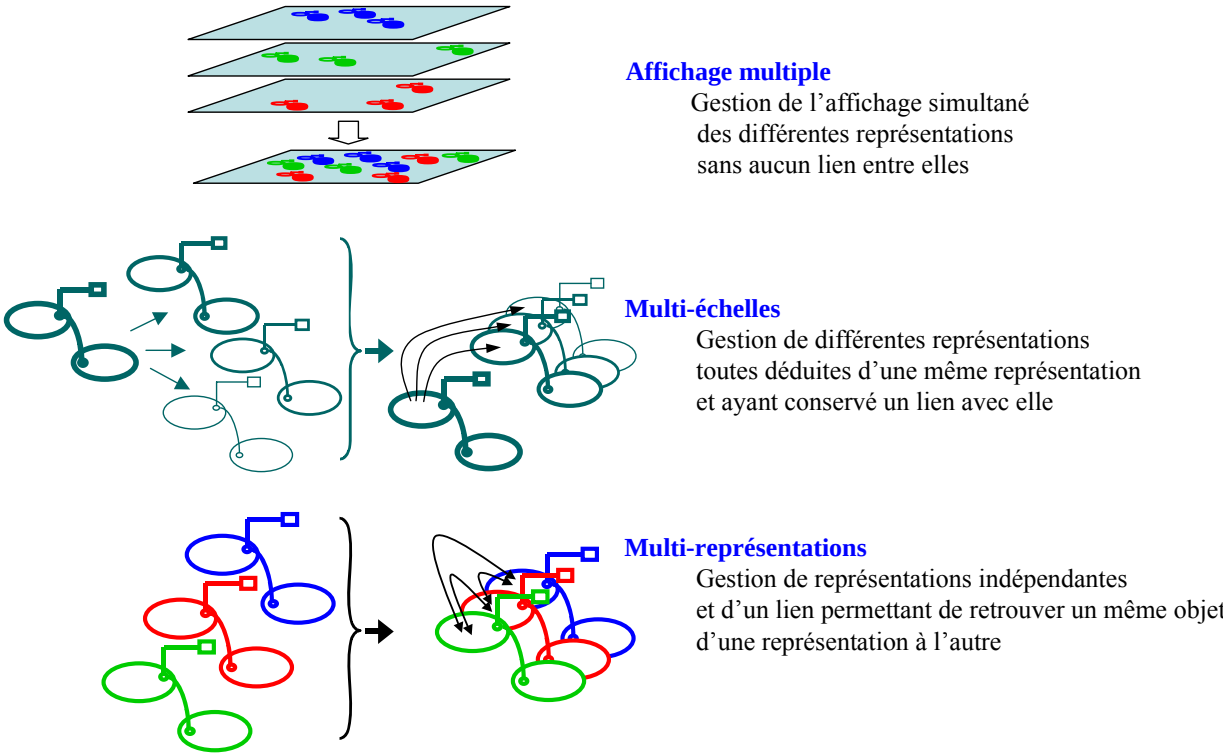

Affichage multiple
Gestion de l'affichage simultané
des différentes représentations
sans aucun lien entre elles
Multi-échelles
Gestion de différentes représentations
toutes déduites d'une même représentation
et ayant conservé un lien avec elle
Multi-représentations
Gestion de représentations indépendantes
et d'un lien permettant de retrouver un même objet
d'une représentation à l'autre

Figure B.2.6. *Différentes façons de gérer la multiplicité des représentations*

B.2.2.2. Solutions pour gérer des représentations multiples

Dans un SIG, si la gestion de l'affichage est standardisée (avec des options propres à certains logiciels), il reste problématique d'établir et de gérer des relations entre des représentations différentes d'un même *objet*, qu'il s'agisse de liens déductifs comme dans le *multi-échelles* (on dérive les bases d'une même base) ou de liens *d'appariement*[5] comme en *multi-représentations*.

5. Dans le A.3. p. 85 nous avons défini l'appariement comme l'opération qui consiste à identifier deux représentations d'un même objet dans deux bases de données différentes

Les enjeux pour les utilisateurs sont cependant importants. En effet ces liens laissent envisager de pouvoir combiner des points de vue différents. Au-delà de la complémentarité des thématiques et de la vision du SIG

fédérateur, on peut envisager la complémentarité des niveaux décisionnels (opérationnel, tactique et stratégique) et entrevoir l'intérêt d'une utilisation des SIG comme outil d'aide à la décision.
Il s'agit toutefois de problèmes techniques. Les éléments de solutions sont donc à puiser dans les recherches des « techniciens de l'information géographique ». On en trouve en particulier auprès des producteurs de données. La gestion de représentations multiples présente en effet pour eux d'autres enjeux : le *multi-échelles* permet d'éviter la constitution et le maintien à jour de bases de données nombreuses [Mustière, 2001]. Et , c'est par la *multi-représentations* qu'ils abordent le thème complexe de la « propagation des mises à jour » d'une base à l'autre et son automatisation [Badard et Lemarié, 2000]. Ils envisagent également d'y recourir pour contrôler la qualité des bases de données et leur cohérence [Sheeren et al, 2004].

B.2.2.1.3. Du multi-niveaux (de détail) vers l'observation et l'analyse multi-niveaux

De notre point de vue, ce qui précède illustre l'évolution des outils SIG. Conçus initialement comme des outils cartographiques, ils tendent depuis à intégrer un spectre toujours plus large d'analyses à référence spatiale et à devenir ainsi des outils de *géomatique* plus performants. La figure B.2.7 montre cette évolution dans la façon de concevoir la multiplicité des représentations dans un SIG.

La première démarche fut effectivement cartographique : les multiples représentations étaient envisagées comme des représentations cartographiques. L'approche était centrée sur la notion d'échelle et intégrait uniquement des *fonctionnalités d'affichage*. La carte devenant base de données, il s'agissait de déduire d'une même base, des données pour des échelles d'observation différentes. On parle d'ailleurs symptomatiquement de *multi-échelles*.

Cependant, la représentation devenant numérique, l'échelle n'a plus suffit à la caractériser et il a fallu introduire les *niveaux de détails*. La représentation devenant plus conceptuelle, elle a par ailleurs utilisé des *fonctionnalités d'abstraction*. On a alors parlé de *multi-représentations* pour signifier qu'il ne s'agissait plus de représenter des informations cartographiques mais plutôt de modéliser des informations géographiques.

L'approche *multi-niveaux* pourrait être l'étape suivante de cette évolution. En effet, en intégrant les *fonctionnalités d'analyses* à une gestion de représentations multiples, elle pourrait opérer le passage d'une observation technique dans un SIG (à partir d'une modélisation multiple) vers une observation plus géographique. Ce faisant, l'approche multi-niveaux réaliserait un glissement sémantique du *multi-niveaux de détails* (qui porte sur l'analyse des spécifications techniques des 'objets') vers le *multiniveau en géographie* (qui porte sur les analyses des *objets*). Pour étudier des phénomènes, l'utilisateur doit pouvoir faire coïncider les deux ; le SIG doit devenir un outil d'analyses à référence spatiale au sens le plus large, où l'analyse spatiale géographique peut exploiter toutes les analyses techniques opérant sur une information géographique.

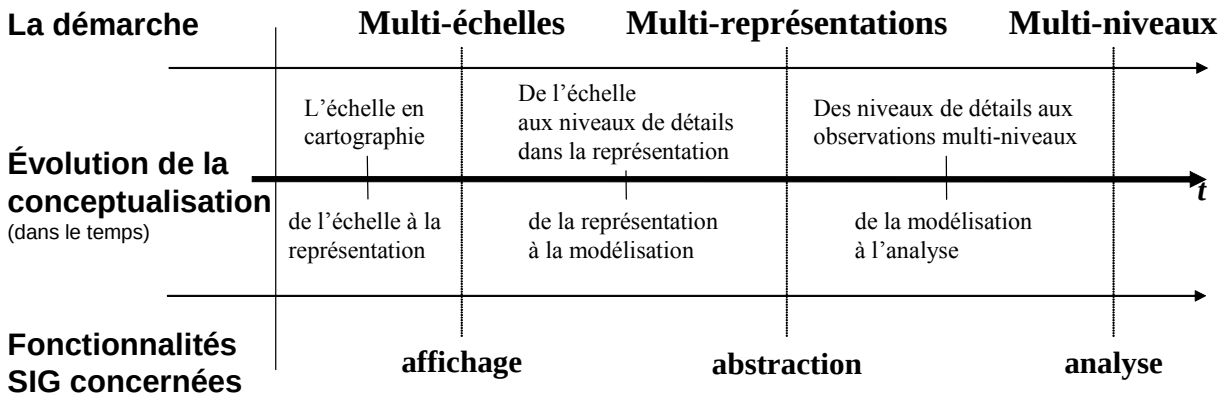


Figure B.2.7. *Une évolution dans la conception de la représentation des informations géographiques : de l'affichage à la modélisation vers l'analyse ; du multi-échelles à la multi-représentations, vers le multi-niveaux*

Partant de l'observation multi-échelles centrée autour de la notion d'échelle d'une carte dans une approche cartographique, il y a eu une évolution vers l'observation multi-représentations centrée non plus sur la « représentation et son l'échelle », mais sur la « représentation comme modélisation » dans une approche plus technique. Cette évolution semble aboutir à l'observation multi-niveaux (de détails) centrée autour des « niveaux de détails pour la modélisation de plusieurs représentations » dans un approche plus analytique.

Le terme de *multi-niveaux* commence d'ailleurs à apparaître pour les SIG. En particulier il a été employé dans le cadre de la généralisation cartographique par [Ruas, 1999], et [Boffet, 2001] l'a mis en œuvre dans une *« méthode de création d'informations multi-niveaux pour la généralisation cartographique de l'urbain »*. Aujourd'hui les démarches *multi-niveaux* dans les SIG ne sont pas très développées, car il reste difficile d'offrir à l'utilisateur les moyens d'analyses conjointes sur plusieurs niveaux de détails. Cela implique en effet, de disposer de liens entre les différentes représentations pour l'affichage, la gestion, la modélisation et aussi pour les analyses elles-même.

Si le *multi-niveaux* peut être considéré comme l'aboutissement d'une évolution technique dans les SIG, il constitue probablement aussi les prémisses d'une autre évolution, plus thématique. La figure B.2.8 montre cette évolution. Partant d'une situation où l'on suppose que les moyens techniques pour une observation multi-niveaux de détails des phénomènes géographiques dans un SIG sont disponibles, alors en s'appuyant sur une démarche d'observation multi-niveaux, les utilisateurs pourront effectuer des analyses multi-niveaux. Les fonctionnalités des SIG profiteront alors aux méthodes de l'analyse multi-niveaux[6] préexistante en géographie, et dans une approche de moins en moins concernée par la composante géométrique et de plus en plus centrée autour de la composante sémantique elles viendront compléter les outils et méthodes d'analyse multiniveau (dans cette orthographe) en statistique. Cette mise en perspective nous permet d'inscrire nos travaux dans une cadre plus large: nous proposons de développer une méthode d'observation *multi-niveaux de détails* dans le SIG adaptée à l'observation multi-niveaux en géographie, comme préambules à des analyses multi-niveaux dans les SIG.

6. comme dans « Analyse spatiale multi-échelles, multi-représentation, analyse multi-niveaux », titre d'un des thèmes proposé pour l'International Conference on Spatial Analysis and GEOmatics, SAGEO'2005.
Toutefois on note que dans [Sanders, 2001] l'orthographe employée est déjà : « multiniveau».

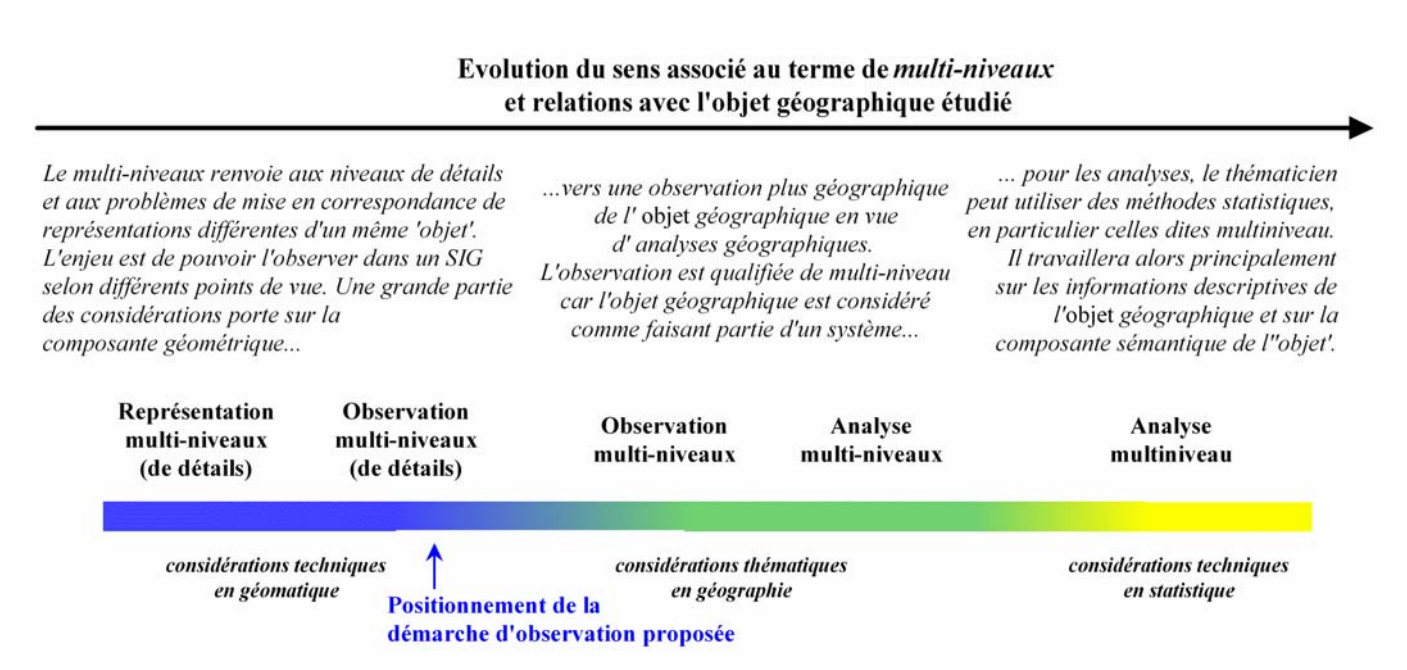

Evolution du sens associé au terme de multi-niveaux
et relations avec l'objet géographique étudié
Le multi-niveaux renvoie aux niveaux de détails et aux problèmes de mise en correspondance de représentations différentes d'un même 'objet'. L'enjeu est de pouvoir l'observer dans un SIG selon différents points de vue. Une grande partie des considérations porte sur la composante géométrique...
...vers une observation plus géographique de l' objet géographique en vue d' analyses géographiques. L'observation est qualifiée de multi-niveau car l'objet géographique est considéré comme faisant partie d'un système...
... pour les analyses, le thématicien peut utiliser des méthodes statistiques, en particulier celles dites multiniveau. Il travaillera alors principalement sur les informations descriptives de l'objet géographique et sur la composante sémantique de l''objet'.
Représentation multi-niveaux (de détails)
Observation multi-niveaux (de détails)
Observation multi-niveaux
Analyse multi-niveaux
Analyse multiniveau
considérations techniques en géomatique
considérations thématiques en géographie
considérations techniques en statistique
Positionnement de la démarche d'observation proposée

Figure B.2.8. *Positionnement d'une méthode d'observation multi-niveaux à partir d'un SIG au sein d'une démarche qui intégrerait différentes étapes d'observations et d'analyses multi-niveaux*

Le schéma montre le glissement sémantique du terme multi-niveaux *entre le sens pris en géomatique dans la «* **représentation** multi-niveaux » *et celui qu'il a en statistiques dans «* l'**analyse** multiniveau ». *Ce glissement se fait lors du passage de* ***l'observation des données*** *géographiques dans le SIG à* ***l'observation des objets*** *en géographie.*

En proposant de mettre au point une méthode d'observation multi-niveaux dans un SIG, nous espérons aider les utilisateurs à analyser les phénomènes qu'ils étudient à différents niveaux et ainsi améliorer les applications SIG. Toutefois, ces travaux visent d'abord l'élaboration d'une méthode pour produire des informations descriptives sur les phénomènes que l'on puisse suivre ensuite simplement dans un SIG.

B.2.3. Vers une méthode de création d'informations multi-niveaux pour l'observation de phénomènes

La méthode de création d'informations multi-niveaux de [Boffet, 2001]

Nous avons vu qu'il existe déjà une *méthode de création d'informations multi-niveaux* (p .124) : il s'agit de la méthode développée par la géographe A. Boffet. Cependant, son objet étant différent du notre, nous avons dû développer une méthode de création d'informations spécifique pour disposer des propriétés qui nous étaient nécessaires. Nous nous situons toutefois dans le prolongement de ces travaux, en particulier :

– nous partageons le constat que : *« La description numérique de l'espace est trop pauvre comparée à la description géographique. Les résultats de traitements sur SIG restent trop éloignés des besoins des utilisateurs en partie parce que les descriptions numériques de l'espace sont souvent peu adaptées ». «Nous avons pour objectif de proposer une méthode de création d'informations adaptées, c'est à dire plus proche des besoins des utilisateurs et de leurs applications »* [Boffet, 2001, p.67].

– nous reprenons l'idée de déduire des informations sur le niveau d'observation principal à partir d'agrégation d'informations à un niveau plus fin,

– nous distinguons trois niveaux d'observation (micro, méso, macro),

La méthode développée dans cette thèse peut présenter des similitudes avec celle de [Boffet, 2001] dans certains cas particuliers où l'objet d'observation principal ne possède pas de représentation préalable sous forme d'information par nature, mais elle s'en distingue en plusieurs points, en particulier :

Méthode de généralisation spécifique / méthode d'observation générique

La méthode de [Boffet, 2001] s'inscrit dans une perspective de « généralisation cartographique » et spécifie le phénomène étudié : « *méthode de création d'informations multi-niveaux, pour la généralisation cartographique de l'urbain* » « *…nous souhaitons créer une information utile à l'automatisation du processus de généralisation cartographique de cartes topographiques* » [Boffet, 2001, p.89]. Il s'agit pour cette méthode de créer des *représentations manquantes* à partir d'une base de données détaillée. Le lien (multi-échelle) entre les différentes représentations découle alors de la construction des représentations (cf. figure B.2.3).

La méthode que nous proposons ne spécifie pas le phénomène (méthode générique) et s'inscrit dans une autre perspective applicative : l'observation de phénomène. En particulier, nous envisageons le fait que l'utilisateur puisse *disposer déjà* d'une représentation sous forme d'information par nature (développée par exemple dans le cadre d'une autre application) et nous nous imposons de l'intégrer alors à l'observation. Cette contrainte assure l'utilisateur de pouvoir observer le phénomène étudié en complémentarité avec les applications préexistantes, sans pour cela reporter sur lui la gestion du problème des représentations multiples (entre la représentation préexistante et celle construite pour l'observation). En revanche, cela implique que le lien entre la représentation sous forme d'agrégation et celle sous forme d'information par nature ne peut pas être uniquement déductif (la représentation par nature pouvant exister indépendamment):

⇒ le lien multi-niveaux ne peut pas être a priori : « est directement déduit de ». Ayant supposé que l'utilisateur puisse travailler par ailleurs avec une représentation de l'objet d'observation principal, il nous faut disposer d'un lien qui puisse s'adapter aux cas où la représentation de l'objet principale pré-existe de façon *indépendante de l'agrégation* des objets élémentaires.

Une unique base de données / possibilité d'intégrer plusieurs types de données

La méthode de [Boffet, 2001] adopte une approche multi-échelle : il s'agit de déduire d'une même base de données (la BD TOPO® en l'occurrence) une représentation pour chaque niveau d'observation.

Nous envisageons la possibilité d'utiliser plusieurs représentations sous forme d'‘agrégation d'objets élémentaires' et d'exploiter pour cela plusieurs types de données éventuellement issues de bases de données d'origines différentes. Par exemple nous envisageons une caractérisation des ‘zones pavillonnaires' à partir des ‘bâtiments', mais aussi des ‘tronçons de route', des ‘parcelles', des ‘ménages', des ‘individus', etc… Il importe en effet pour nous de développer une méthode ayant la capacité de construire de *nombreuses* informations à observer[8]. Nous espérons ainsi donner les moyens d'une observation plus fine, d'un suivi plus détaillé (avec en perspective l'idée de saisir des nuances utiles à l'identification de phénomènes émergents et/ou de bifurcations). Or le choix de la méthode d'agrégation influe sur le fait que les agrégations puissent servir à

8. L' « Analyse à référence spatiale » a été définie dans ce sens : il s'agit de pouvoir mobiliser l'ensemble des moyens (outils et méthodes) disponibles capables de créer des informations à partir de données géographiques.

caractériser la même observation : par exemple, l'utilisation de buffers pour agréger les 'objets élémentaires' conduit généralement à des représentations différentes de l''objet' d'observation principal sous forme d'agrégation (cf. figure B.2.9). On se trouve alors à nouveau confronté au problème des représentations multiples.

⇒ Souhaitant mettre en commun les caractéristiques calculées à partir d''agrégations' différentes, la construction des 'agrégations' doit donc conduire à une même représentation de l'objet d'observation principal sous forme d'agrégation et ce quelque soit les 'objets élémentaires' considérés.

Prise en compte de la dimension temporelle

Enfin, [Boffet, 2001] n'aborde pas l'aspect temporel de l'observation qui est au cœur de cette thèse.

⇒ La méthode développée dans ces travaux doit permettre un suivi commode des 'objets' observés.

Pour une méthode de création d'informations multi-niveaux pour l'observation de phénomènes géographiques

Pour construire une observation à plusieurs niveaux de détails dans un SIG, nous somme partis du modèle à trois niveaux d'observation (micro, méso, macro) de [Boffet, 2001, p.70] (repris de [Ruas, 1999], employé également par [Bard, 2004]) :

« Le niveau micro décrit les objets géographiques de précision sémantique et spatiale maximale, ce qui correspond à la granularité des objets initiaux ».

« Le niveau méso est le niveau intermédiaire, défini par une composante spatiale et une composante sémantique. Un objet méso est composé d'objets micro ou méso. Il peut posséder sa propre sémantique. Par exemple un quartier est un objet méso composé de bâtiments et de rues. Un groupe de maisons est un autre objet méso sans posséder une dénomination particulière autre que « groupe de bâtiments ».

« Le niveau macro permet de décrire les connaissances et les contraintes sur les populations d'informations, telles que toutes les routes ou tous les bâtiments. Un objet macro est défini par composition sémantique d'objets micro et méso » [Boffet, 2001, p.70].

Nous les avons cependant redéfinis pour les adapter à notre besoin. Il s'agit pour nous de construire des informations sur le niveau méso à partir d'agrégation de données sur le niveau micro, mais nous considérons pour cela que l''objet' *méso* est prescrit : il préexiste, qu'il soit ou non dans la base de données, et nous cherchons comment le qualifier à partir d''objets' *micro* (fig. B.2.10). Si nous élaborons les caractéristiques *méso* dans une démarche ascendante des *objets micro* vers *l'objet méso*, notre premier mouvement va d'abord du *méso* vers le *micro* ; pas seulement dans une réflexion sur *« comment allons-nous le construire ? »*, mais surtout *« à partir de quelles informations, de quelles données, de quel existant allons nous le caractériser ? »*. C'est là que se rejoue pour nous la problématique de la multi-représentations : dans le lien à réaliser entre la représentation par nature et la représentation comme agrégation. Ainsi, dans ces travaux :

– le *niveau méso* correspond au niveau d'observation du phénomène : celui qui permet de l'identifier simplement, de le situer rapidement, celui où l''objet' d'observation principal est représenté sous forme d'information par nature.

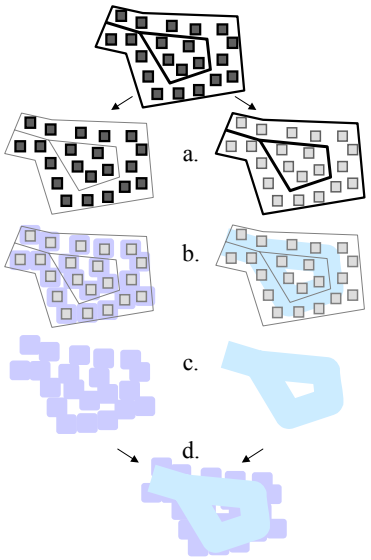


Figure B.2.9. *Lorsque l'on considère des 'objets élémentaires' distincts, l'agrégation par buffer peut conduire à des représentations sous forme d''agrégation d'objets élémentaires' différentes, qu'il est alors difficile d'étudier conjointement.*

Par exemple, si on considère sur une même zone de référence les objets élémentaires 'bâtiments' (colonne de gauche) et 'tronçons de route' (colonne de droite) (a), la construction des buffers élémentaires (b) conduira à des buffers agrégés de formes différentes (c) dont les composantes géométriques seront différentes et ne pourront donc pas servir à décrire le « même » 'objet' d'observation principal. (d).

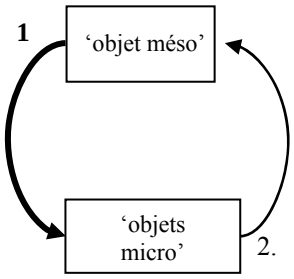


Figure B.2.10. *L''objet méso' préexiste aux 'objets micro' qui sont utilisés pour le caractériser*

– le *niveau micro* correspond dans une première approche à l'ensemble des informations géographiques de granularité minimale, et plus spécifiquement au niveau des informations susceptibles de contribuer à une agrégation qui ait du sens pour le phénomène.

– Adoptant une approche technique par *thèmes* (cf. Annexe 2), [Ruas,1999, p.73] définit le niveau *macro* comme le niveau regroupant l'ensemble des 'objets' de même type (ex. : l'ensemble des 'bâtiments' ou encore l'ensemble des 'tronçons de routes'). Pour nous, le niveau *macro* prend un sens géographique plus général, mais aussi moins précis. Nos travaux exploitent peu ce niveau. Il sert essentiellement de cadre à l'analyse pour saisir le contexte du phénomène étudié. Il désigne suivant les cas :

- soit une information géographique de granularité thématique plus large (ex. : la 'ville', l''agglomération urbaine', la 'région') qui correspond à une approche thématique spécialisée,
- soit un territoire géographique regroupant l'ensembles des informations géographiques disponibles sur la zone étudiée, appréhendé au travers de la 'zone *technique* d'étude' (autrement dit, la 'zone' couverte par les jeux de données utilisées).

Cette définition plus générale permet de ne pas restreindre les champs thématiques de l'observation[9].

9. Nous imaginons cependant que l'observation géographique puisse s'élargir en terme de territoire et diminuer en niveaux de détails et que les 'objets' macro puissent devenir des 'objets' méso, dans une analyse emboîtée du type de celle présentées par [Boffet, 2001]. Dans ce cas nous préciserions les 'objets' macro observés avant de les désigner comme 'objets' méso et de définir un nouveau niveau macro.

Pour définir la méthode de création d'informations multi-niveaux, il nous faut établir un lien entre la représentation comme information par nature de l'*objet* méso et les représentations comme agrégation de cet *objet*, qui soit simple à exploiter dans les SIG. Nous proposons de développer pour cela une méthode d'agrégation spécifique : la **méthode d'agrégation par emprise**.

B.2.4. Méthode d'agrégation par emprise dans un SIG

B.2.4.1. Les méthodes d'agrégation

En nous appuyant sur l'expérience des géographes, nous avons montré dans le paragraphe précédent l'intérêt d'envisager une observation à plusieurs niveaux pour étudier un phénomène. Une fois désigné l'*objet* servant à caractériser le phénomène, nous avons proposé de prendre en compte des *objets* supplémentaires, à un niveau plus fin, pour élaborer des informations descriptives additionnelles. Dans ce paragraphe, nous supposerons avoir identifié préalablement :

– l'*objet* d'observation principal utilisé pour caractériser le phénomène ; il définit le *niveau méso*. Nous appellerons 'objet méso', sa représentation sous forme d'information géographique par nature (surfacique).

– et des *objets* élémentaires pertinents vis à vis du phénomène, c'est à dire dont l'agrégation puisse servir à construire des informations complémentaires ; ils définiront le *niveau micro*. Leur représentation sous forme d'information par nature sera alors appelée 'objet micro'.

Par exemple pour observer l'*urbanisation pavillonnaire* nous avons choisi comme *objet* d'observation principal les *zones pavillonnaires*, et comme *objets* d'analyse principaux les *bâtiments*. Les 'zones pavillonnaires' seront alors les 'objets méso', et les 'bâtiments' des 'objets micro'.

Il s'agit ici de développer une méthode pour construire une 'agrégation d'objets micro', de façon à ce que les informations caractérisant cette 'agrégation' puissent être utilisées facilement pour enrichir la description de l''objet méso' (fig. B.2.11).

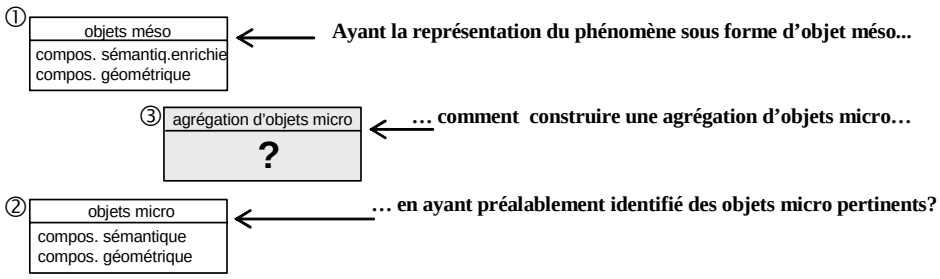


Figure B.2.11. *La problématique de l'agrégation ou comment construire l'agrégation des objets micro ?*

Le niveau méso et le niveau micro étant respectivement déterminés par la représentation sous forme d'information par nature de l'objet d'observation principal (dit 'objet méso') ① et par la représentation des objets élémentaires (dits 'objets micro')②, il s'agit ici de développer une méthode pour construire une 'agrégation d'objets micro'③ permettant d'enrichir la caractérisation de l''objet méso'

B.2.4.2. Caractérisation des méthodes d'agrégations d'objets élémentaires

Ayant un ensemble d''objets micro' (fig. B.2.12), plusieurs types d'agrégation peuvent être envisagés qui aboutissent généralement à des représentations variées. Des logiques diverses les différencient ainsi que les techniques mises en œuvre. Par exemple :

- dans la figure B.2.13a., la logique de l'agrégation est sémantique. Ayant un ensemble d''objets', on les agrège par classe, en fonction de leurs attributs caractéristiques. On utilise pour cela des méthodes statistiques.

- dans la figure B.2.13b., la logique de l'agrégation est une logique de proximité spatiale. Ayant un ensemble d''objets', on construit pour chacun une 'zone' (définie par exemple comme son voisinage immédiat, son aire d'influence, etc.) en utilisant les fonctionnalités des SIG dites de buffer (ou zone tampon). La taille des buffers est généralement paramétrée en fonction du problème posé. Les 'zones élémentaires' sont ensuite agrégées et tous les 'objets' appartenant à une 'zone agrégée' connexe sont alors regroupés.

- dans la figure B.2.13c., la logique de l'agrégation est aussi spatiale. Ayant un ensemble d''objets' d'un certain type, on définit leur enveloppe convexe à l'aide de méthodes mathématiques. On agrège alors l'ensemble des 'objets' situés à l'intérieur de cette enveloppe. On peut ensuite décider de suivre ou bien le contenu de cette enveloppe, ou bien son emprise.

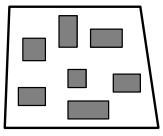


Figure B.2.12. 'Objets micro' à agréger

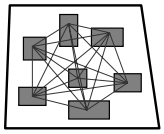


Figure B.2.13a. *L'agrégation par classification*

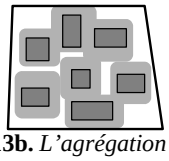


Figure B.2.13b. *L'agrégation par buffers*

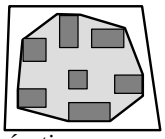


Figure B.2.13c. *L'agrégation par enveloppe*

Ces trois méthodes d'agrégations ont en commun de partir des 'objets micro' pour construire l''agrégation' (fig. B.2.14) :

- en a. on part des attributs caractéristiques des 'objets',
- en b. on part des 'objets' pour construire les buffers élémentaires,
- en c. on part des 'objets' pour construire leur enveloppe convexe.

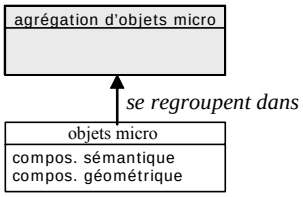


Figure B.2.14. *Agrégation partant des 'objets micro'*

En revanche, ces trois méthodes se différencient par la forme du résultat :

- en a., le résultat de l'agrégation fournit des sous-ensembles d''objets', la méthode servant à les sélectionner,
- en c., le résultat de l'agrégation est fourni sous la forme d'une règle d'appartenance à une surface spécifiée (ici l'enveloppe convexe),
- Le cas b. est hybride. La méthode par buffers agrégés peut conduire soit directement à un sous-ensemble d''objets', ceux dont l'agrégation des buffers forme une surface convexe, soit, en référence à cette surface, à l'ensemble des 'objets' qui y sont inclus. C'est par exemple cette deuxième forme qui est considérée dans la figure B.2.3, pour « construire » la *ville* à partir de l''agrégation des bâtiments'.

B.2.4.3. Méthode d'agrégation par emprise

Dans le cadre de ces travaux nous introduisons une quatrième méthode. La logique de celle-ci est une logique spatiale proche de la méthode c. Comme elle, son résultat se présente sous forme d'une règle d'appartenance à une surface (fig. B.2.15). Cependant, la définition de la surface de référence n'est pas la même. Il ne s'agit pas de l'enveloppe convexe, ni de toute autre surface construite à partir des 'objets' micro (méthode b.). Dans la méthode que nous envisageons, **l'agrégation est** fondamentalement géométrique : Elle est réalisée sur la **base d'une emprise au sol déterminée de façon indépendante des 'objets micro'**. De façon imagée elle fonctionne à la façon d'un emporte-pièce, où la forme étant donnée, on sélectionne (extrait) ce qui est à l'intérieur (fig. B.2.16). Par la suite nous la désignerons sous l'appellation *méthode d'agrégation par emprise*.

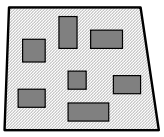


Figure B.2.15. *L'agrégation spatiale par emprise*

Figure B.2.16. *Principe de l'emporte-pièce*

On notera ici, que dans un emporte-pièce, *ce qui est donné n'est pas une surface mais un contour, et celui-ci permet ensuite de sélectionner une surface. Dans la méthode d'agrégation proposée, nous partons indifféremment d'un contour ou d'une surface, car dans le SIG "l'emporte-pièce" sera de toute façon traduit sous forme d'une surface. En effet, la plupart des SIG " raisonnant" sur des « objets à l'intérieur de surface » et non sur des « objets à l'intérieur d'un contour », il est nécessaire de construire au préalable la surface que le contour délimite.*

Cette agrégation possède une propriété importante pour nous : elle assure la *compatibilité* entre 'agrégations' d'objets micro différents, la **compatibilité** voulant dire ici que l'on peut trouver des représentations en informations géographiques qui aient la même composante géométrique support, ce qui permet de les étudier conjointement. En effet, lorsque l'on applique la méthode d'agrégation par emprise à deux types d''objets micro' distincts, on obtient bien deux 'agrégations' compatibles, leur support géométrique commun étant alors la composante géométrique servant à représenter la 'portion de territoire' utilisée comme emprise de référence. Par exemple, si l'on considère comme dans la figure B.2.17a l''emprise' d'une 'zone de référence' (représentée par le polygone en hachures) et deux ensembles d''objets micro' différents, alors la figure B.2.17b. montre comment l'emprise en hachures permet de sélectionner indépendamment deux sous-ensembles d''objets micro' (les 'bâtiments' dans la première colonne et les 'tronçons de route' dans la seconde), auxquels on peut associer une composante géométrique commune, celle de l''emprise' ayant servis à les agréger.

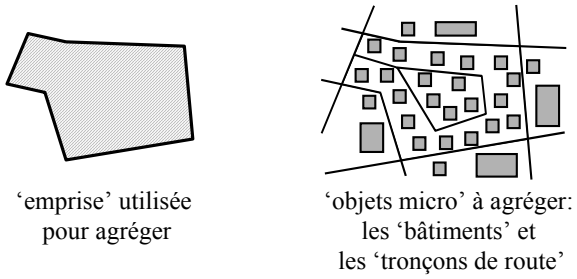


Figure B.2.17a. *'Emprise de référence' et 'objets' micro à agréger*

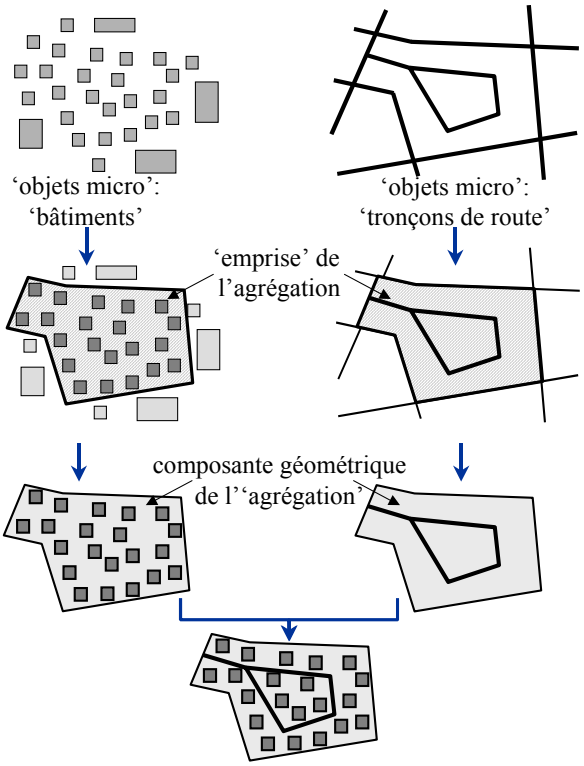


Figure B.2.17b. *La compatibilité entre agrégations élémentaires par la méthode par emprise*

Toutes les méthodes d'agrégations ne possèdent pas cette propriété. Comme le montre la figure B.2.8 (reprise de la p. 127): si l'on utilise parallèlement la méthode des buffers pour agréger deux types d''objets micro' sur une même 'portion de territoire', alors l'emprise des deux 'agrégations' obtenues ne seront pas *compatibles*. Par exemple si on considère les 'objets micro' 'bâtiments' (colonne de gauche) et 'tronçons de route' (colonne de droite) (a), la construction des buffers élémentaires (b) conduira à des buffers agrégés d'emprises différentes (c) qui ne seront pas *compatibles* (d).

Dans la *méthode par emprise*, la surface d'agrégation est déterminée de façon indépendante des 'objets micro'. Il s'agit donc de la définir. Nous allons maintenant montrer qu'un choix judicieux de cette surface de référence peut permettre de dépasser les contraintes (les limites) que posent habituellement la multi-représentations dans les SIG.

Si l'on considère l''objet méso', en tant que représentation sous forme d'information par nature (surfacique), il possède une composante géométrique de type polygone, pour caractériser son *emprise au sol*. Nous proposons alors d'adopter précisément *cette* 'emprise au sol' comme surface de référence et *ce* polygone comme emporte-pièce pour la méthode d'agrégation. (Par exemple dans la figure B.2.18, colonne de gauche, nous agrégerons les 'bâtiments'② selon l''emprise de la zone pavillonnaire'① représentée en hachures).

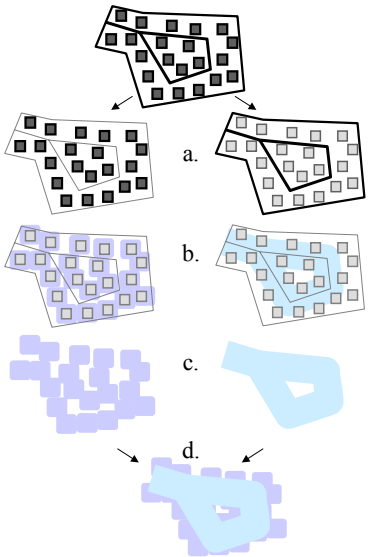


Figure B.2.8. *Lorsque l'on considère des 'objets élémentaires' distincts, l'agrégation par buffer peut conduire à des représentations sous forme d'"agrégation d'objets micro' non compatibles*

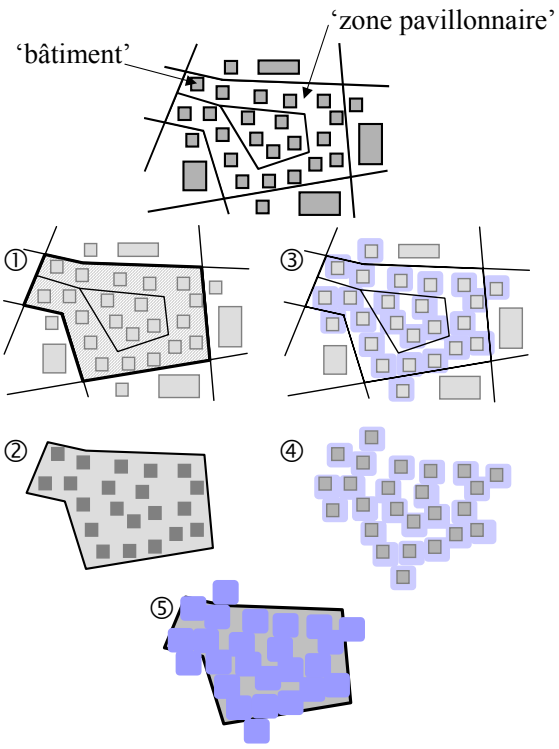


Figure B.2.18. *Compatibilité entre la représentation de l'objet méso et les 'agrégations'*

Nous avons vu précédemment que le choix de *la méthode d'agrégation par emprise* nous assure la compatibilité entre 'agréations' d'objets micro différents. En désignant spécifiquement l''emprise de l'objet méso' comme surface d'agrégation de référence, nous nous assurons en plus la *compatibilité* avec la représentation de *l'objet méso*. Ce qu'illustre la gauche de la figure B.2.18, où considérant une 'zone pavillonnaire' dont l'emprise (en gris) sert à définir l'emprise d'agrégation (en hachure) utilisée pour agréger les 'bâtiments' situés dans cette 'zone' ①, on obtient en ② deux 'objets' (une 'agrégation de bâtiments' et l'objet 'zone pavillonnaire) qui ont par construction une emprise commune, ce qui n'aurait pas été le cas avec une agrégation par buffers. En effet, comme le montre la droite de la figure, l'agrégation par buffers des 'bâtiments' ③ possède une emprise ④ qui n'est pas *compatible* avec celle de la 'zone pavillonnaire', leur composante géométriques sont différentes ⑤ l'agrégation par buffer intégrant même des 'bâtiments' en dehors de la 'zone'.

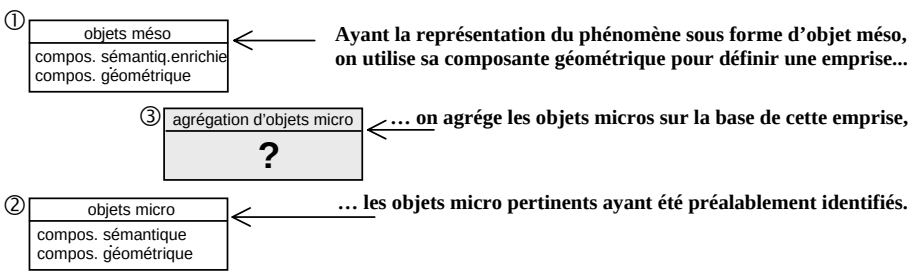

①
objets méso
compos. sémantiq.enrichie
compos. géométrique
Ayant la représentation du phénomène sous forme d'objet méso, on utilise sa composante géométrique pour définir une emprise...
③
agrégation d'objets micro
?
… on agrége les objets micros sur la base de cette emprise,
②
objets micro
compos. sémantique
compos. géométrique
… les objets micro pertinents ayant été préalablement identifiés.

Figure B.2.19. *La méthode d'agrégation par emprise pour surmonter le problème des représentations multiples*

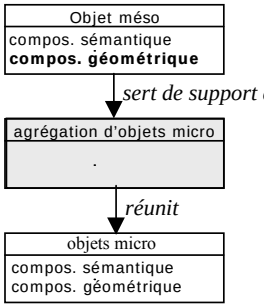


Figure B.2.20. *Construction de l''agrégation' par emprise à partir de la composante géométrique de l''objet*

B.2.4.4. La méthode d'agrégation par emprise pour une méthode d'observation multi-niveaux dans les SIG

Pour effectuer une observation multi-niveaux dans un SIG, nous proposons de construire les 'agrégations d'objets micro' en utilisant une méthode d'agrégation spécifique aux SIG (car fondée sur l'exploitation de la composante géométrique et sur la représentation sous forme d'information par nature) : la méthode d'agrégation par emprise. Nous proposons d'agréger les 'objets micro' sur la base d'une *emprise au sol*, cette *emprise* étant choisie égale à celle définie par la composante géométrique de l''objet méso' (cf. figure B.2.19 et B.2.20). La surface d'agrégation étant indépendante des 'objets micro', nous nous assurons ainsi la *compatibilité* des 'agrégations' entre elles. Cette première propriété nous permet d'envisager l'exploitation d''objets micro' de types différents (par exemple pour caractériser les 'zones pavillonnaires' on peut utiliser les 'bâtiments', mais aussi 'les tronçons de routes', les 'parcelles', etc. cf. figure B.2. ?) et ce faisant de pouvoir construire plus d'informations complémentaires. Par ailleurs, le choix particulier de la surface d'agrégation nous garantit la *compatibilité* entre la représentation de l''objet méso' et les 'agrégations d'objets micro', puisque par construction, ils partagent la même composante géométrique (fig. B.2.21).

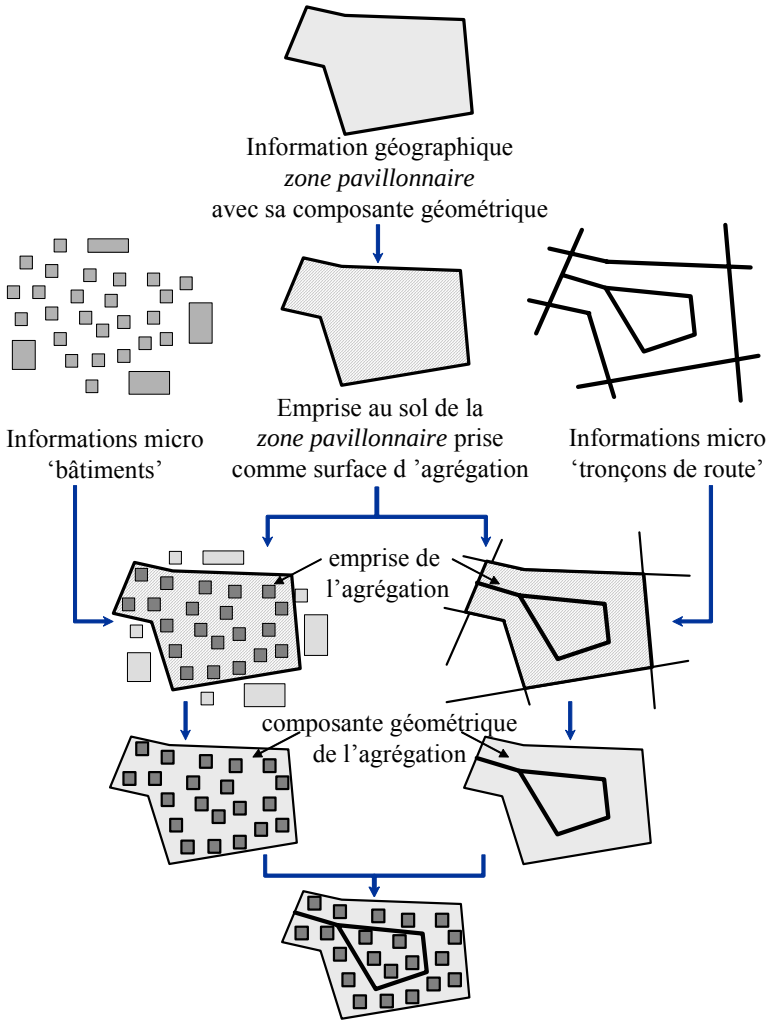
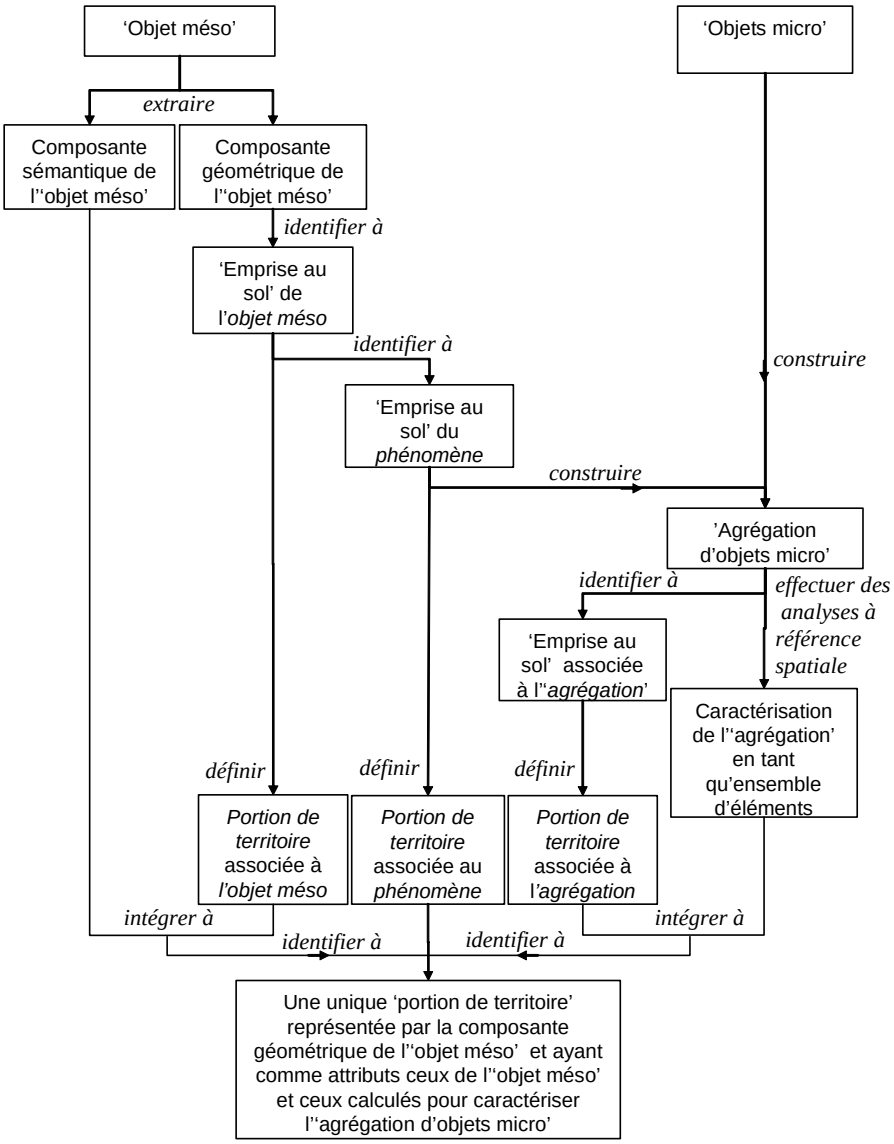


Figure B.2.21. *La méthode d'agrégation par emprise assure la compatibilité des représentations, ce qui leur assure de pouvoir faire référence à une même composante géométrique*

Cette deuxième propriété de partager la même composante géométrique établit un lien entre les deux types de représentations analysées dans le SIG, celle de l''objet méso' et celle des 'agrégations d'objets micro'. Ce lien peut se traduire formellement par : les *agrégations d'objets micro* déterminées par cette méthode et l'*objet méso* partagent la même *emprise au sol ;* elles ont pour support géographique une même *portion de territoire*. *Objet méso* et *agrégation d'objets micro* peuvent alors être considérés comme des caractéristiques de cette *portion de territoire* et leurs caractéristiques peuvent alors être transféré formellement à la *portion de territoire*. Ainsi dans la pratique, la méthode d'agrégation par emprise permet d'utiliser la même composante géométrique, pour servir de support à la fois aux caractéristiques de l''objet méso' et aux caractéristiques des 'agrégations des objets micro'

B.2.5. Méthode d'observation multi-niveaux dans un SIG

B.2.5.1. Principe de la méthode d'observation multi-niveaux


'Objet méso'
'Objets micro'
extraire
Composante sémantique de l''objet méso'
Composante géométrique de l''objet méso'
identifier à
'Emprise au sol' de l'objet méso
identifier à
construire
'Emprise au sol' du phénomène
construire
'Agrégation d'objets micro'
identifier à
effectuer des analyses à référence spatiale
'Emprise au sol' associée à l''agrégation'
Caractérisation de l''agrégation' en tant qu'ensemble d'éléments
définir
définir
définir
Portion de territoire associée à l'objet méso
Portion de territoire associée au phénomène
Portion de territoire associée à l'agrégation
intégrer à
intégrer à
identifier à
identifier à
Une unique 'portion de territoire' représentée par la composante géométrique de l''objet méso' et ayant comme attributs ceux de l''objet méso' et ceux calculés pour caractériser l''agrégation d'objets micro'

Ayant une représentation de l''objet méso', on extrait sa composante géométrique.

Cette composante caractérise l'*emprise au sol* de l'*objet méso*. Elle peut donc formellement être identifiée à l''emprise au sol de l'*objet méso'*.

Ayant identifiée l'*emprise au sol* de l'*objet méso* à celle du *phénomène*, la composante géométrique peut alors également être identifiée à l''emprise au sol du phénomène'.

Si on utilise cette composante pour définir une méthode d'agrégation par emprise que l'on applique ensuite aux 'objets micro', on obtient alors par construction une 'agrégation d'objets micro' qui à pour *emprise au sol* celle du phénomène.

A l'aide des fonctionnalités d'analyses à référence spatiale, on caractérise l''agrégation des objets micro' en tant qu'ensemble d'objets. On peut ainsi obtenir de nombreux indicateurs sur cet ensemble.

Ces indicateurs qui décrivent l'ensemble des *objets micro* situés sur la *portion de territoire* définie par *l'emprise de l'agrégation*, peuvent être considéré comme des caractéristiques de cette *portion de territoire*.

Par ailleurs les attributs de l''objet méso' peuvent être de la même façon considérés comme des caractéristiques de la 'portion de territoire' qu'il occupe.

Or l''objet méso', le 'phénomène' et l'agrégation d'objets micro' possédant par construction la même composante géométrique (du fait de la propriété de compatibilité assurée par de la méthode d'agrégation utilisée), il en résulte que les trois *portions de territoire* correspondant à leur *emprise* respective sont identiques. On peut donc les identifier à une unique emprise décrite alors par la composante géométrique de l''objet méso' et ayant pour composante sémantique l'ensemble des attributs décrivant l''objet méso' et l''agrégation d'objets micro'

Figure B.2.22. *le principe de la méthode d'observation multi-niveaux dans un SIG*

Le schéma de la figure B.2.22 illustre le cheminement formel qui permet d'effectuer pratiquement le transfert des attributs de l''objet méso' et des attributs calculés pour caractériser l''agrégation d'objets micro' vers la même 'portion de territoire' (à la même composante géométrique)[10]. Il justifie ainsi l'utilisation de la méthode d'agrégation par emprise pour mettre en œuvre une méthode d'observation multi-niveaux ; en effet comme le montre ce schéma, la méthode d'agrégation par emprise permet de construire sans développement supplémentaire des 'agrégations d'objets micro' observables conjointement avec la représentation de l'*objet méso*.

10. en jargon informatique on appelle cela : « une descente des attributs sur la géométrie ».

Si on simplifie le schéma général de mise en œuvre en ôtant les étapes purement formelles (fig. B.2.23), nous obtenons le schéma pratique de la méthode (fig. B.2.24).

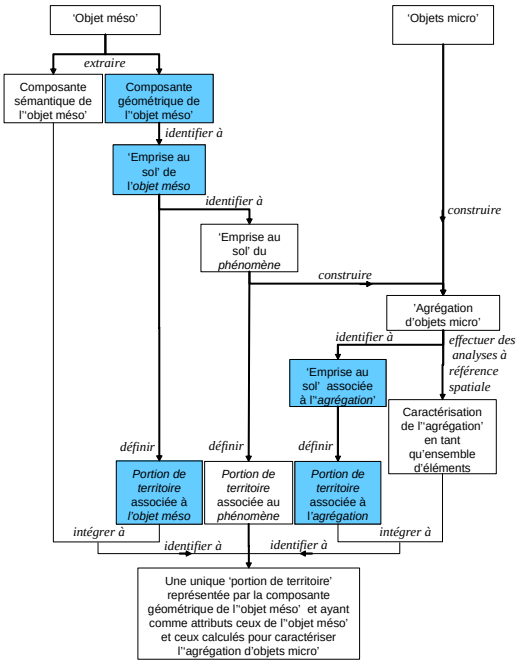

'Objet méso'
'Objets micro'
extraire
Composante sémantique de l''objet méso'
Composante géométrique de l''objet méso'
identifier à
'Emprise au sol' de l'objet méso
identifier à
construire
'Emprise au sol' du phénomène
construire
'Agrégation d'objets micro'
identifier à
effectuer des analyses à référence spatiale
'Emprise au sol' associée à l'agrégation'
Caractérisation de l''agrégation' en tant qu'ensemble d'éléments
définir
définir
définir
Portion de territoire associée à l'objet méso
Portion de territoire associée au phénomène
Portion de territoire associée à l'agrégation
intégrer à
identifier à
identifier à
intégrer à
Une unique 'portion de territoire' représentée par la composante géométrique de l''objet méso' et ayant comme attributs ceux de l''objet méso' et ceux calculés pour caractériser l''agrégation d'objets micro'

Figure B.2.23. *Les étapes purement formelles de la méthode d'observation multi-niveaux*

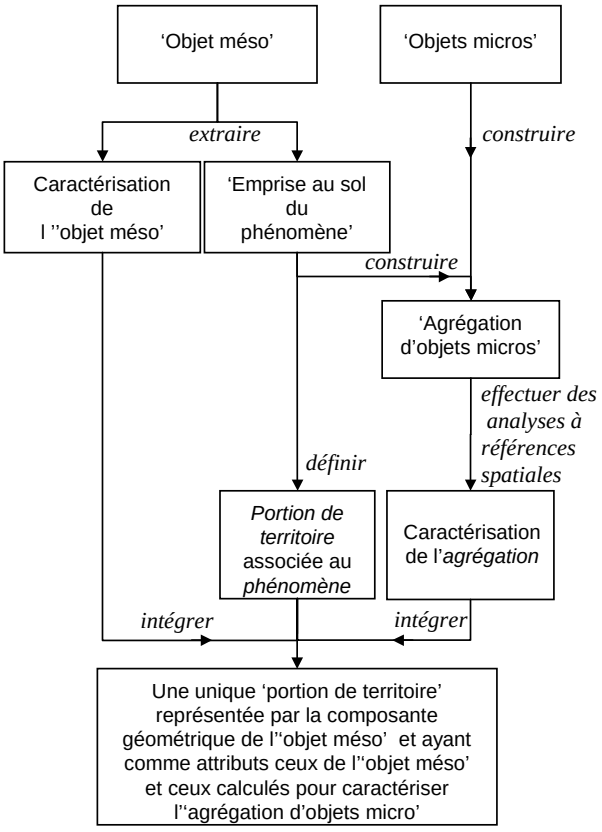

'Objet méso'
'Objets micros'
extraire
construire
Caractérisation de l ''objet méso'
'Emprise au sol du phénomène'
construire
'Agrégation d'objets micros'
effectuer des analyses à références spatiales
définir
Portion de territoire associée au phénomène
Caractérisation de l'agrégation
intégrer
intégrer
Une unique 'portion de territoire' représentée par la composante géométrique de l''objet méso' et ayant comme attributs ceux de l''objet méso' et ceux calculés pour caractériser l''agrégation d'objets micro'

Figure B.2.24. *Schéma pratique de la méthode multi-niveaux : les différentes étapes de sa mise en oeuvre*

Toutefois, notons ici qu'une condition essentielle pour pouvoir appliquer ce schéma de mise en œuvre de la méthode proposée, et donc de pouvoir effectuer une observation multi-niveaux, repose sur une hypothèse de départ :

« l'utilisateur dispose d'une représentation de l'objet méso ».

B.2.5.2. Cas où l'emprise d'agrégation n'est pas donnée

Pour présenter le principe de la méthode d'observation multi-niveaux nous nous sommes placés dans le cas où l'on disposait d'une représentation de l''objet méso' et nous avons montré comment enrichir cette description en introduisant des informations obtenues à partir d'un niveau plus fin. Il nous faut cependant maintenant envisager les situations où le thématicien désigne comme *objet d'observation principal* adéquat un *objet* pour lequel il n'existe pas de représentation sous forme d'information par nature (fig. B.2.25 et B.2.26) (voir Partie C, le cas de l'*ambiance urbaine*). Dans le cas où cette représentation n'existe pas, l'utilisateur ne dispose pas non plus de la composante géométrique associée pour lui servir d'« emporte-pièce » pour construire les 'agrégations d'objets micro'. En revanche, puisqu'il dispose de données à un niveau inférieur dont il sait qu'elles peuvent servir à construire des indicateurs intéressants, la question se pose alors pour lui de savoir comment constituer un 'objet' au niveau méso qui ait du sens pour le phénomène : qui puisse être identifié à l'emprise du phénomène et servir à agréger les 'objets micro' pour le caractériser. Cet 'objet' qui sert à mettre en œuvre la méthode d'agrégation par emprise joue un rôle essentiellement technique : il assure de pouvoir mettre en relation des représentations d''agrégations d'objets micro' différentes (relation de compatibilité) et de disposer d'une représentation sous forme d'information par nature simple à observer (et par la suite à suivre) dans le SIG. Nous avons choisi d'appeler cet objet technique *objet méso-structurel*.

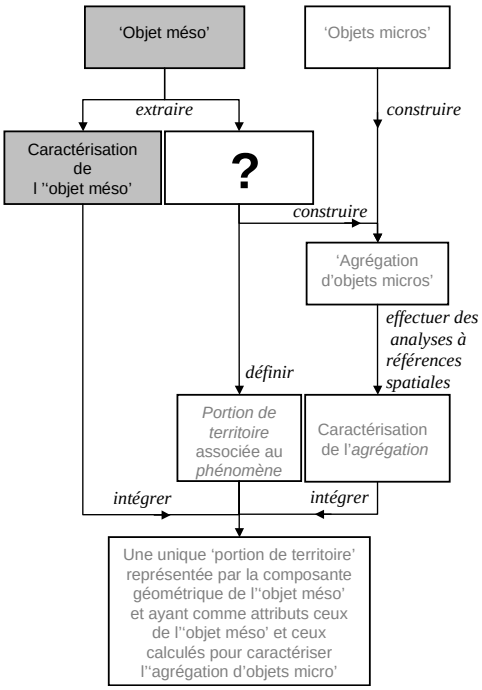
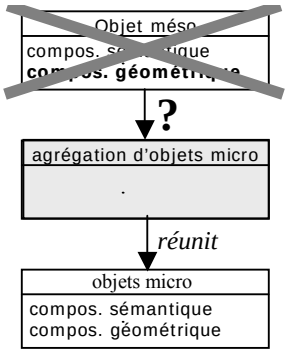
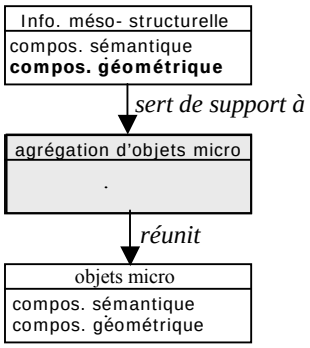

'Objet méso'
'Objets micros'
extraire
construire
Caractérisation de l ''objet méso'
?
construire
'Agrégation d'objets micros'
effectuer des analyses à références spatiales
définir
Portion de territoire associée au phénomène
Caractérisation de l'agrégation
intégrer
intégrer
Une unique 'portion de territoire' représentée par la composante géométrique de l''objet méso' et ayant comme attributs ceux de l''objet méso' et ceux calculés pour caractériser l''agrégation d'objets micro'

Figure B.2.25. *Lorsqu'il n'y a pas d''objet méso'*



Quand on ne possède pas de représentation sous forme d'objet méso, comment définir le support d'agrégation?



L'information méso-structurelle permet d'étendre la méthode d'agrégation aux cas où l'on ne possède pas de représentation sous forme d'objet méso

Figure B.2.26. *Le problème du manque d''objet méso' : introduction de l'information méso-structurelle*

B.2.5.2.1. Introduction du concept d'information méso-structurelle

Le concept d'information méso-structurelle permet d'étendre la méthode d'observation multi-niveaux ; il nous affranchit de la contrainte portant sur l'existence de données pour représenter l'*objet* d'observation principal sous forme d'information géographique par nature.

Une information méso-structurelle est un objet méthodologique conçu dans le SIG ; définie par sa composante géométrique, elle correspond à un objet géographique 'portion de territoire'. Elle est utilisée pour définir l''emprise au sol du phénomène' et sert de support pour l'agrégation des 'objets micro' (fig. B.2.27). Une information est dite *information méso-structurelle* pour un phénomène, si son *emprise* est compatible avec celle de *l'objet* désigné comme objet adéquat d'observation pour le phénomène, *compatible* au sens où sa composante géométrique pourrait caractériser l''emprise' de l''objet méso' ou bien si on peut trouver d'autres informations de même type telles que la réunion de leur composante géométrique forment une partition de l''emprise' du phénomène. (Par exemple, pour observer l'urbanisation pavillonnaire, on peut choisir comme information méso-structurelle une 'zone du POS' dont l'emprise correspondra à celle de la 'zone pavillonnaire' désignée comme objet d'observation adéquat, ou un îlot' qui participe d'un ensemble dont la réunion aura pour emprise celle de la 'zone pavillonnaire').

⇒ Autrement dit l'information méso-structurelle s'identifie à l'information géographique 'portion de territoire associée à l'emprise du phénomène' ou à l'une de ses parties qui ait du sens par rapport au phénomène étudié.

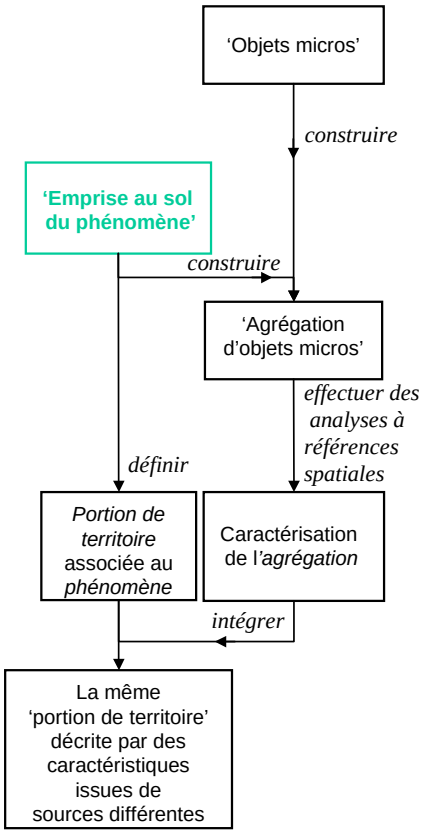


Figure B.2.27. *Lorsqu'il n'y a pas d''objet méso'…*

B.2.5.2.2. Utilisation du concept d'information méso-structurelle

Lorsque l'on ne possède pas de données directes pour décrire le phénomène, il peut être intéressant de choisir pour *information méso-structurelle* des informations géographiques existantes (ex. : un 'îlot' pour l'urbanisation pavillonnaire, une 'zone industrielle' pour l'imperméabilisation des sols, voir Partie C) qui même si elles ne sont pas directement liées à l'*objet d'observation* souhaité, peuvent correspondre à un niveau d'observation compatible. L'aide d'un thématicien est requise pour déterminer des *objets pertinents* pour le phénomène étudié ; leurs caractéristiques propres peuvent alors servir à caractériser le phénomène. Par exemple, si on s'intéresse à une 'zone pavillonnaire', mais que l'on ne dispose pas de données sur cet 'objet', on peut choisir d'utiliser des 'îlots' pour définir son 'emprise' et utiliser ensuite les caractéristiques de ces *îlots* décrivant leurs *habitants* pour caractériser la 'zone pavillonnaire'. Ainsi, une *même* information géographique peut avoir *plusieurs fonctions au sein de cette méthode* : être une *information géographique* directement exploitable pour qualifier les *objets* auxquels on s'intéresse et servir d'information méso-structurelle pour déterminer l''emprise au sol'.

Vu l'enjeu de l'information méso-structurelle, il nous paraît important de spécifier quel rôle certaines informations géographiques tiennent au sein de la méthode d'observation :

- nous les qualifierons d'*informations méso-structurelles* lorsqu'elles serviront uniquement en tant que 'portion de territoire', pour définir le mode d'agrégation grâce à leur composante géométrique,

- nous les spécifierons par le terme d'*informations méso-géographiques* lorsque nous nous intéresserons à elles à la fois pour leur composante géométrique et pour leur composante sémantique. Les *informations méso-géographiques* se réfèrent à un 'objet géographique' mieux spécifié que la 'portion de territoire' (ex. : 'zone industrielle' pour l'étude des *zones imperméables*).

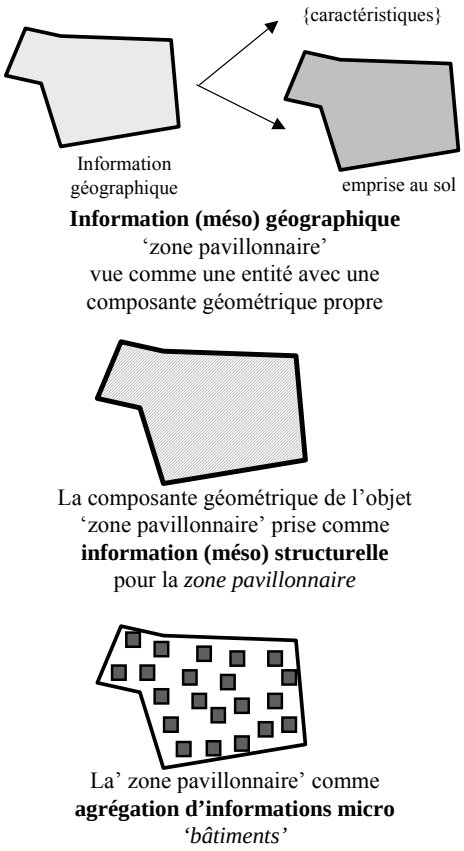

{caractéristiques}
Information géographique
emprise au sol
Information (méso) géographique
'zone pavillonnaire'
vue comme une entité avec une composante géométrique propre
La composante géométrique de l'objet 'zone pavillonnaire' prise comme information (méso) structurelle pour la zone pavillonnaire
La' zone pavillonnaire' comme agrégation d'informations micro 'bâtiments'

Figure B.2.28. *Information méso-géographique/ méso-structurelle*

Par exemple, les informations géographiques 'zones pavillonnaires' sont méso-géographiques lorsqu'on s'intéresse à la superficie des *zones pavillonnaires*. Elles sont méso-structurelles lorsque leur composante géométrique est utilisée pour agréger un ensemble de 'pavillons' (fig. B.2.28). La 'zone industrielle' est une information méso-géographique quand on recherche la date de création de la *zone industrielle* et une information méso-structurelle quand on utilise sa composante géométrique pour agréger les 'zones imperméables élémentaires'.

Il importe que l'information méso-structurelle soit appréhendée au niveau méso indépendamment des 'objets micro'. Sa composante géométrique doit permettre d'agréger les 'objets micro', mais elle doit éviter de résulter d'une telle agrégation pour pouvoir assurer la compatibilité avec les autres 'agrégations' d'objets micro possibles.

B.2.5.2.3. Types d'informations méso-structurelles suivant les conditions de l'application

Si nous avons choisi d'attribuer un terme spécifique à chaque fonction de l'information au sein de la méthode d'observation, c'est pour lui permettre d'être générique ; la spécification des fonctions permet en effet de décomposer le processus et de mieux identifier les problèmes à résoudre, et donc par la suite les moyens à mettre en œuvre pour le faire. Certes il existe des cas simples où une même information peut servir à la fois pour :

- la localisation du phénomène,
- l'agrégation des informations élémentaires
- l'enrichissement des informations descriptives,

(Par exemple, l''enceinte industrielle' qui existe dans la BD TOPO® sous forme d'information par nature surfacique, peut servir à :

- la *localisation* directe et à l'observation du phénomène de développement économique,
- l'agrégation des 'entreprises',
- l'enrichissement des informations descriptives en apportant ses propres caractéristiques (ex. : superficie) pour décrire le phénomène.

Dans ces cas particuliers, l'observation statique du phénomène étudié ne soulève alors pas de difficultés avec un SIG. Cependant, la méthode doit permettre de répondre à des cas plus complexes, où chacune de ces fonctions peut faire appel à une 'information' différente, parfois inexistante. Nous distinguons cinq situations (fig. B.2.29) :

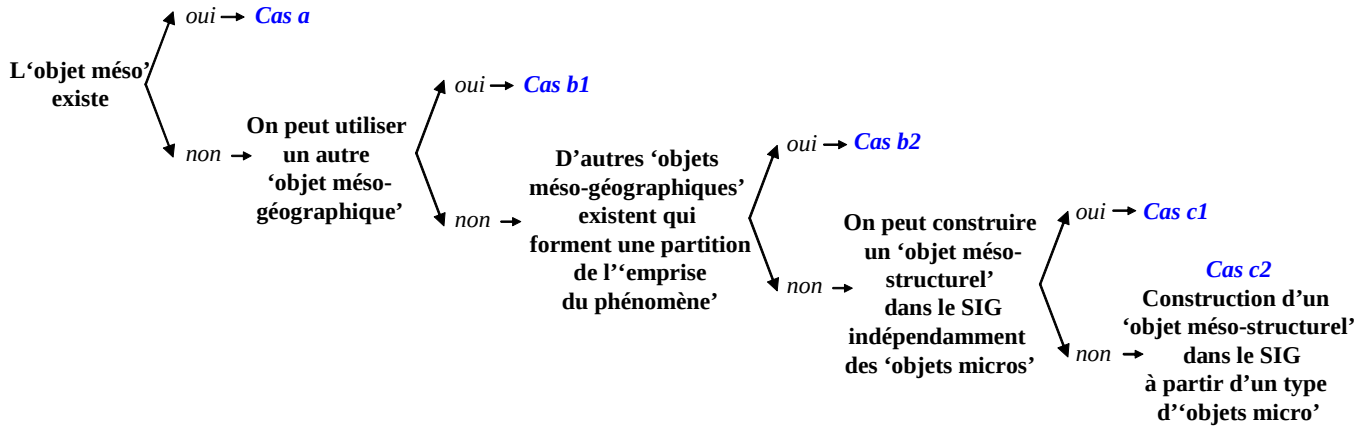

L'objet méso' existe
oui → Cas a
non → On peut utiliser un autre 'objet méso-géographique'
oui → Cas b1
non → D'autres 'objets méso-géographiques' existent qui forment une partition de l''emprise du phénomène'
oui → Cas b2
non → On peut construire un 'objet méso-structurel' dans le SIG indépendamment des 'objets micros'
oui → Cas c1
non → Cas c2 Construction d'un 'objet méso-structurel' dans le SIG à partir d'un type d''objets micro'

Figure B.2.29. *Différents cas pour définir l'information méso-structurelle*

a- Il existe une représentation de l'*objet d'observation* principal sous forme d'information par nature, 'objet méso'.

Ce cas est le plus simple (fig. B.2.30) ; c'est celui qui nous a servi à expliquer le principe général de la méthode. L'information méso-géographique, 'objet méso' sert également d'information méso-structurelle : sa composante géométrique devient un support « naturel » pour agréger les 'objets micro' utilisés pour caractériser le phénomène.

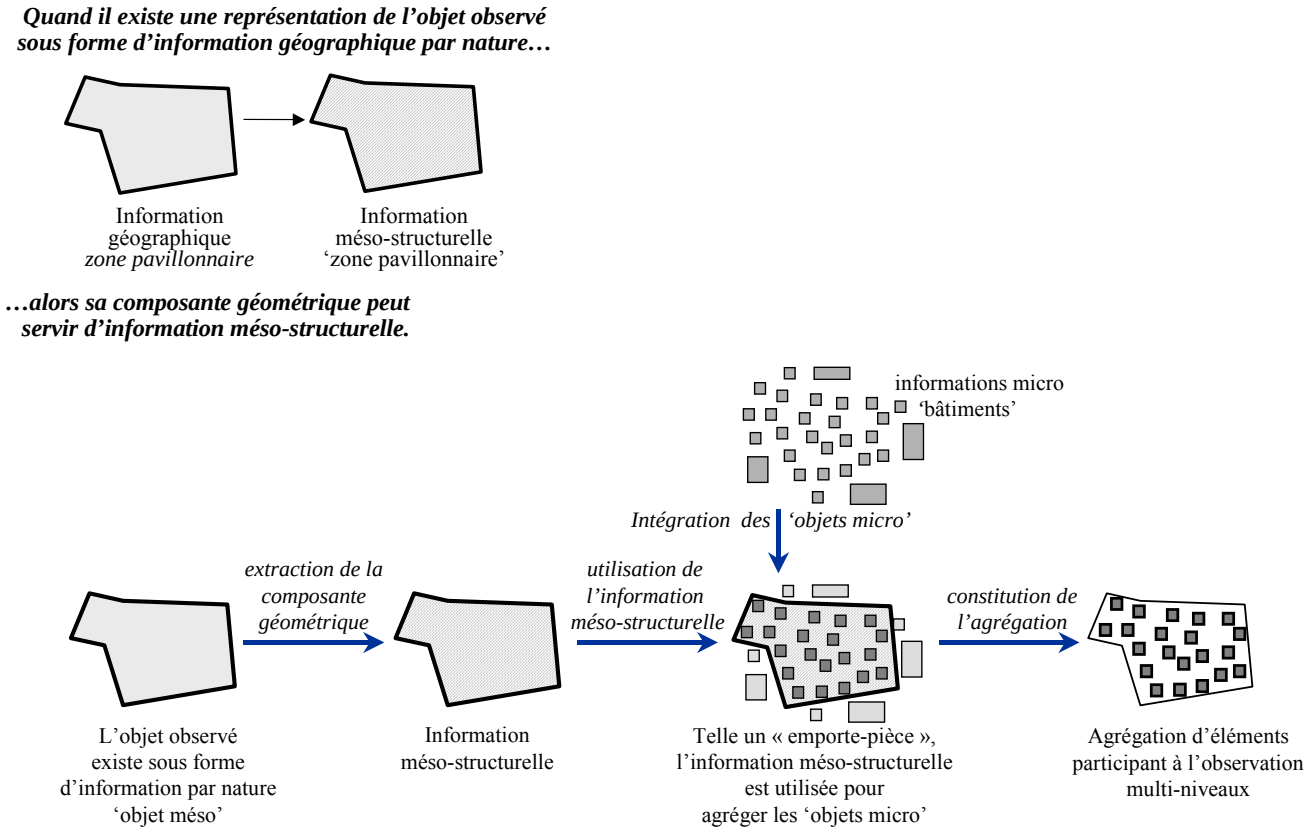

Quand il existe une représentation de l'objet observé sous forme d'information géographique par nature…
Information géographique zone pavillonnaire
Information méso-structurelle 'zone pavillonnaire'
…alors sa composante géométrique peut servir d'information méso-structurelle.
informations micro 'bâtiments'
Intégration des 'objets micro'
extraction de la composante géométrique
utilisation de l'information méso-structurelle
constitution de l'agrégation
L'objet observé existe sous forme d'information par nature 'objet méso'
Information méso-structurelle
Telle un « emporte-pièce », l'information méso-structurelle est utilisée pour agréger les 'objets micro'
Agrégation d'éléments participant à l'observation multi-niveaux

Figure B.2.30. *Méthode d'observation multi-niveaux quand l''objet méso' existe*

b- La représentation de 'l'objet méso' n'existe pas.

b1- On peut parfois trouver dans les bases de données un autre *objet* géographique pour jouer le rôle d'information méso-structurelle.

N'ayant pas de représentation sous forme d'information par nature, on ne dispose pas non plus de la composante géométrique nécessaire pour agréger les 'objets micro'. Dans certains cas cependant, avec l'aide d'un thématicien, on peut trouver un *objet géographique* qui possède une *emprise* compatible avec l'*objet d'observation principal* et qui dispose d'une représentation sous forme d'information par nature. La *compatibilité* des *emprises* permet alors, après identification des 'emprises', d'utiliser la composante géométrique de ces nouveaux 'objets' pour construire les 'agrégations d'objets micro' utiles à la caractérisation du phénomène étudié (fig. B.2.31). Par exemple, l'emprise de certains *quartiers* correspond à celle d'une *zone pavillonnaire*. Le 'quartier' peut alors servir d'information méso-structurelle en remplacement de la 'zone pavillonnaire' pour construire les 'agrégations d'objets micro' servant à caractériser l'urbanisation pavillonnaire.

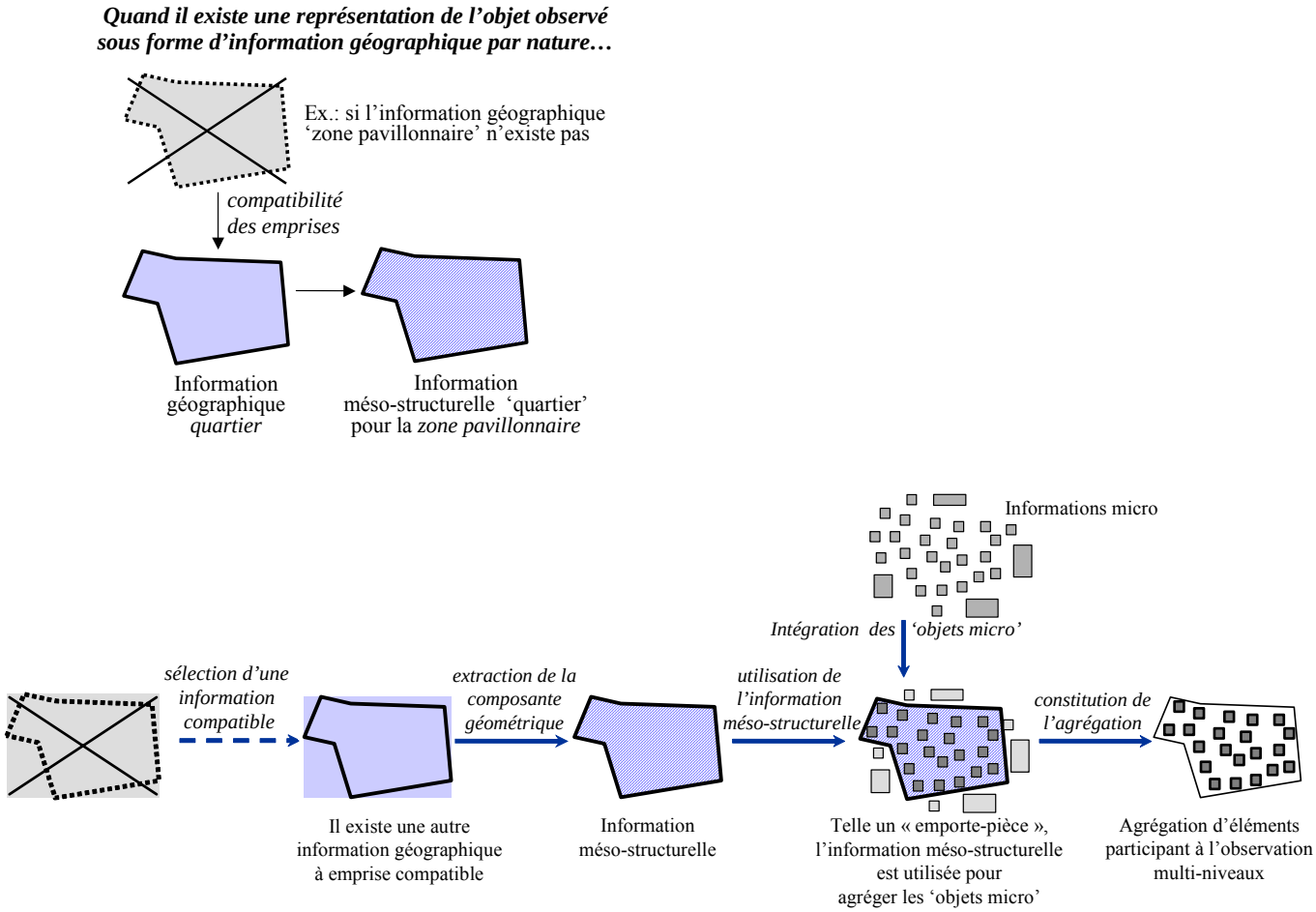
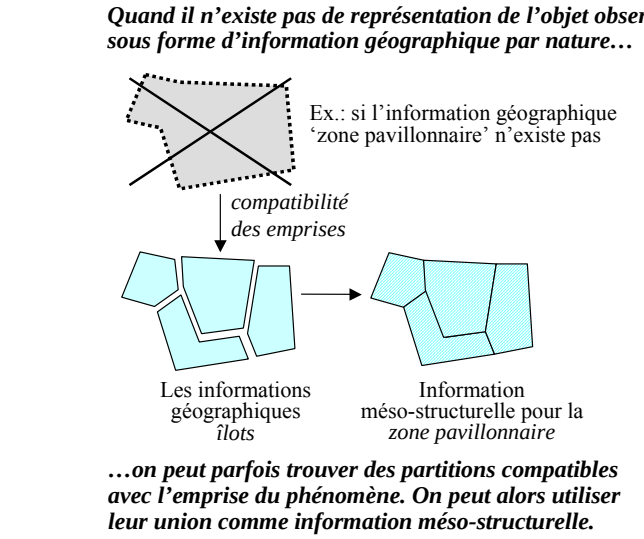

Quand il existe une représentation de l'objet observé sous forme d'information géographique par nature…
Ex.: si l'information géographique 'zone pavillonnaire' n'existe pas
compatibilité des emprises
Information géographique quartier
Information méso-structurelle 'quartier' pour la zone pavillonnaire
Informations micro
Intégration des 'objets micro'
sélection d'une information compatible
extraction de la composante géométrique
utilisation de l'information méso-structurelle
constitution de l'agrégation
Il existe une autre information géographique à emprise compatible
Information méso-structurelle
Telle un « emporte-pièce », l'information méso-structurelle est utilisée pour agréger les 'objets micro'
Agrégation d'éléments participant à l'observation multi-niveaux

Figure B.2.31. *Méthode d'observation quand il existe une information méso-structurelle de substitution*

b2- Parfois, il ne suffit pas de prendre *un* objet géographique ; il en faut plusieurs.

Par exemple, si la 'zone pavillonnaire' n'existe pas, on peut effectuer l'agrégation des 'bâtiments' sur la base des 'parcelles' ou bien des 'îlots' qui la constituent (fig. B.2.32).


Quand il n'existe pas de représentation de l'objet observé sous forme d'information géographique par nature…
Ex.: si l'information géographique 'zone pavillonnaire' n'existe pas
compatibilité des emprises
Les informations géographiques îlots
Information méso-structurelle pour la zone pavillonnaire
…on peut parfois trouver des partitions compatibles avec l'emprise du phénomène. On peut alors utiliser leur union comme information méso-structurelle.

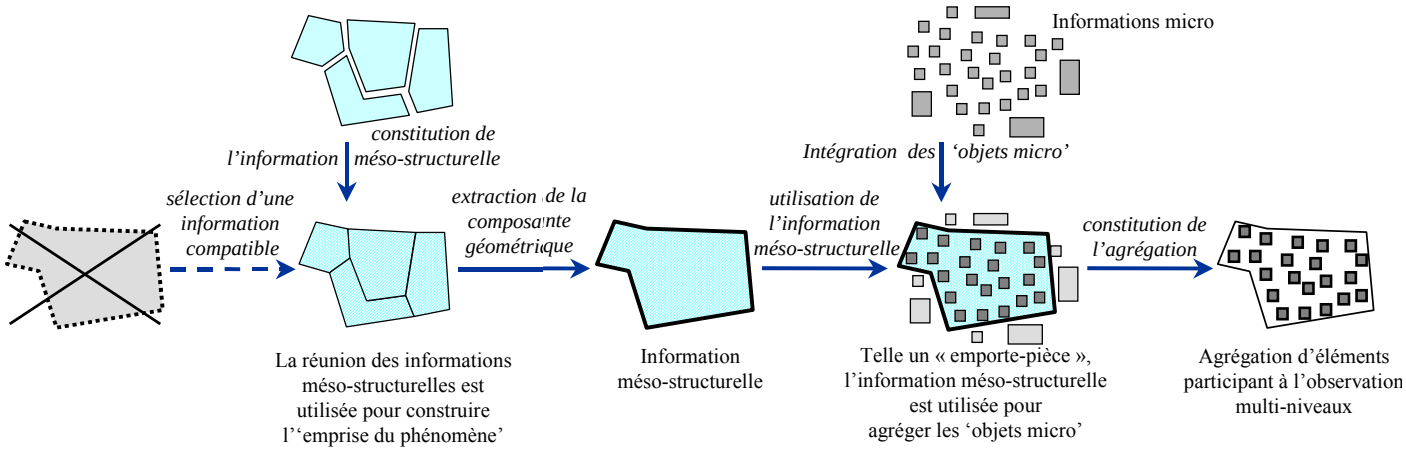

Informations micro
constitution de l'information méso-structurelle
Intégration des 'objets micro'
sélection d'une information compatible
extraction de la composante géométrique
utilisation de l'information méso-structurelle
constitution de l'agrégation
La réunion des informations méso-structurelles est utilisée pour construire l''emprise du phénomène'
Information méso-structurelle
Telle un « emporte-pièce », l'information méso-structurelle est utilisée pour agréger les 'objets micro'
Agrégation d'éléments participant à l'observation multi-niveaux

Figure B.2.32. *Méthode d'observation quand il n'existe pas d'information méso-structurelle couvrant la totalité de l''emprise du phénomène'*

c- Il n'existe pas d''objet méso-géographique' pouvant servir à construire l''emprise du phénomène'.

Il faut donc construire une information méso-structurelle pour agréger les 'objets micro' et que cette information ait du sens pour le phénomène étudié.

c1- Dans le cas où l'on ne parvient pas à trouver d''objet géographique' pour jouer le rôle d'information méso-structurelle, il s'agit d'en construire un. Pour préserver la compatibilité entre les 'agrégations', il importe de continuer à concevoir cet 'objet' au niveau méso, c'est à dire *indépendamment* des 'objets micro' (cf. figure B.2.8). Pour conserver à la méthode sa simplicité, nous proposons de résoudre ces cas en utilisant des buffers au niveau méso.

En identifiant l'‘emprise du phénomène' à un « voisinage » d'‘objet méso' que l'on peut construire sous la forme d'un buffer, on obtient une information méso-structurelle pour agréger les ‘objets micro' (fig. B.2.33). Par exemple, dans le cas de l'application au suivi de l'*Ambiance urbaine* (cf. chapitre C), nous avons choisi d'associer l'ambiance aux *tronçons de rue*. Le ‘tronçon de la rue' est une information par nature, mais de type polyligne ; nous ne pouvions pas utiliser sa composante géométrique comme une « surface » d'agrégation. Nous avons donc défini un voisinage de type « buffer » autour du ‘tronçon de rue' pour servir d'‘emprise d'agrégation'.

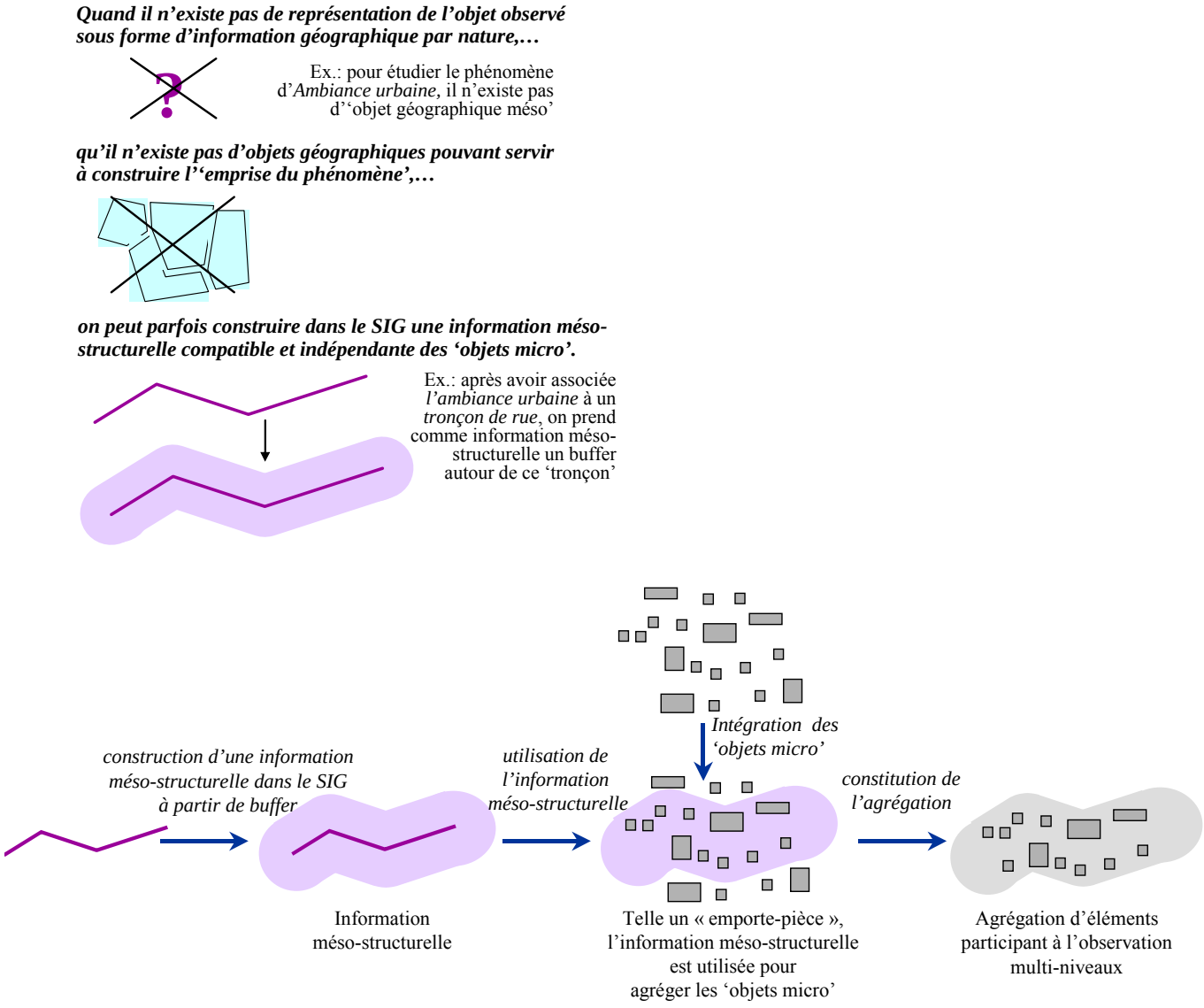
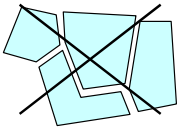

Quand il n'existe pas de représentation de l'objet observé sous forme d'information géographique par nature,…
Ex.: pour étudier le phénomène d'Ambiance urbaine, il n'existe pas d'‘objet géographique méso'
qu'il n'existe pas d'objets géographiques pouvant servir à construire l'‘emprise du phénomène',…
on peut parfois construire dans le SIG une information méso-structurelle compatible et indépendante des ‘objets micro'.
Ex.: après avoir associée l'ambiance urbaine à un tronçon de rue, on prend comme information méso-structurelle un buffer autour de ce ‘tronçon'
Intégration des ‘objets micro'
construction d'une information méso-structurelle dans le SIG à partir de buffer
utilisation de l'information méso-structurelle
constitution de l'agrégation
Information méso-structurelle
Telle un « emporte-pièce », l'information méso-structurelle est utilisée pour agréger les ‘objets micro'
Agrégation d'éléments participant à l'observation multi-niveaux

Figure B.2.33. *Méthode d'observation quand il n'existe pas d'information méso-structurelle*

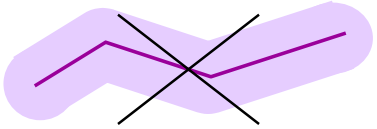
c2- La construction des ‘agrégations' *à partir* des ‘objets micro' sera réservée aux cas extrêmes, où aucune des autres solutions ne peut être appliquée. En effet, cette agrégation demande plus d'expertises techniques et thématiques (définir la bonne stratégie d'agrégation : pour la mettre en œuvre, pour définir les conditions et les paramètres de l'agrégation, etc.). Par ailleurs, si l'on peut envisager prendre ‘l'emprise de l'agrégation' comme ‘emprise au sol', la compatibilité avec d'éventuelles ‘agrégations' additionnelles risque d'être difficile à obtenir (cf. figure B.2.34).

Quand il n'existe pas de représentation de l'objet observé sous forme d'information par nature,…

qu'il n'existe pas d''objet géographique' permettant de construire l' emprise du phénomène',…



que l'on ne parvient pas à construire d'information géographique méso-structurelle par buffer,…



on a peu de moyens pour observer les phénomènes. On pourra alors chercher à construire des agrégations à partir d''objets micro'.

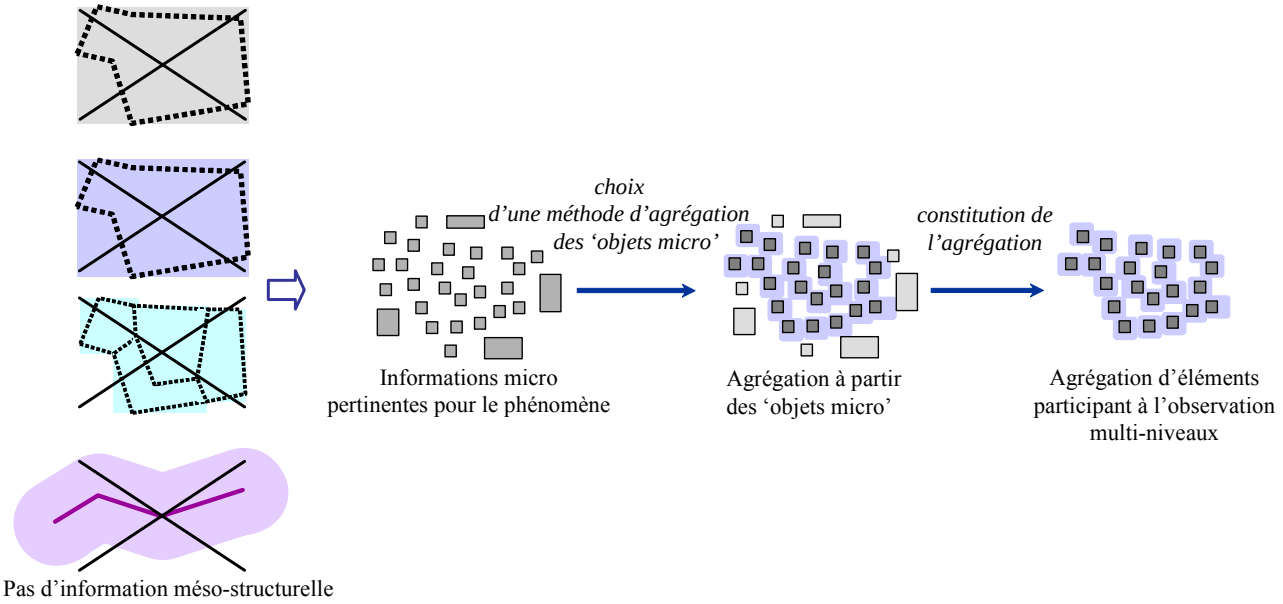


Figure B.2.34. *Méthode d'observation quand il n'existe pas d'information méso-géographique structurelle*

Lorsque l'on ne parvient pas à trouver ou à construire une information méso-structurelle au niveau méso (c'est à dire qui soit indépendante des 'objets micro') pour observer le phénomène étudié, il faut non seulement trouver des informations micro pertinentes vis à vis de ce phénomène, mais également définir une stratégie d'agrégation en fonction : du phénomène, des objets micro agrégés et éventuellement des autres 'objets micro' susceptibles d'être agrégés.

Ainsi dans le cas de la figure, on constate que l'on obtient une 'emprise' très différentes de celle attendue, qui peut même avoir des trous. Pour éviter les trous et récupérer les caractéristiques propres seulement à un type d''objet micro', on peut utiliser une enveloppe, appliquer des fonctions de lissage et de filtrage, nécessitant alors de savoir choisir les bonnes fonctions, les bonnes valeurs de paramètres, etc.,.

Remarque : Cet dernier type d''agrégation' est en revanche très bien adapté lorsque l'observation ne porte pas en priorité sur une description riche et détaillée du phénomène étudié, mais sur une observation directe des évolutions de *l'emprise du phénomène*. Il convient ainsi pour certaines études sur des phénomènes diffus dont l'emprise est évolutive (par exemple pour suivre *l'évolution de l'emprise* d'une épidémie).

La figure B.2.35 présente de façon synthétique le **schéma de la méthode d'observation multi-niveaux**. Plus général que le schéma présenté tout d'abord (qui ne correspondait qu'au cas où il existe un 'objet méso', voir figure B.2.24), il permet d'intégrer les différents cas rencontrés par les utilisateurs :

- en noir, nous avons extrait la partie centrale et générique de la méthode : la construction d''agrégation d'objets micro' sur la base d'une 'emprise' pour pouvoir dans un second temps par analyses à références spatiales, en déduire des caractéristiques à associer à cette *emprise*,

- en bleu-vert, nous avons représenté l'étape préliminaire indispensable : la détermination de l'information méso-structurelle nécessaire à l'agrégation.

La figure B.2.36 en double page suivante, résume les différentes conditions et options d'application de cette méthode.

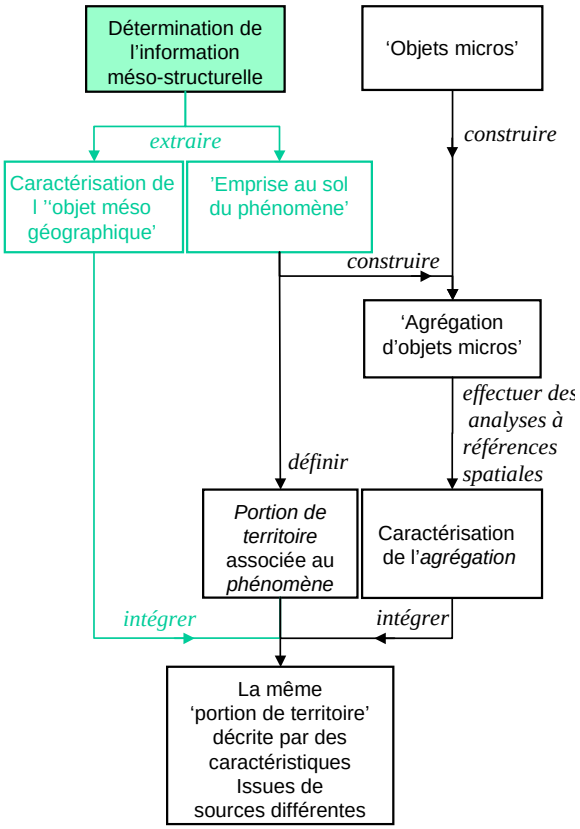


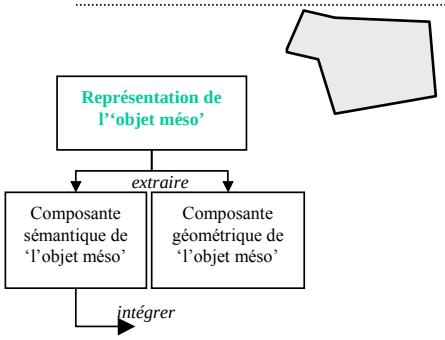
Figure B.2.35. *Schéma général de la méthode d'observation multi-niveaux*

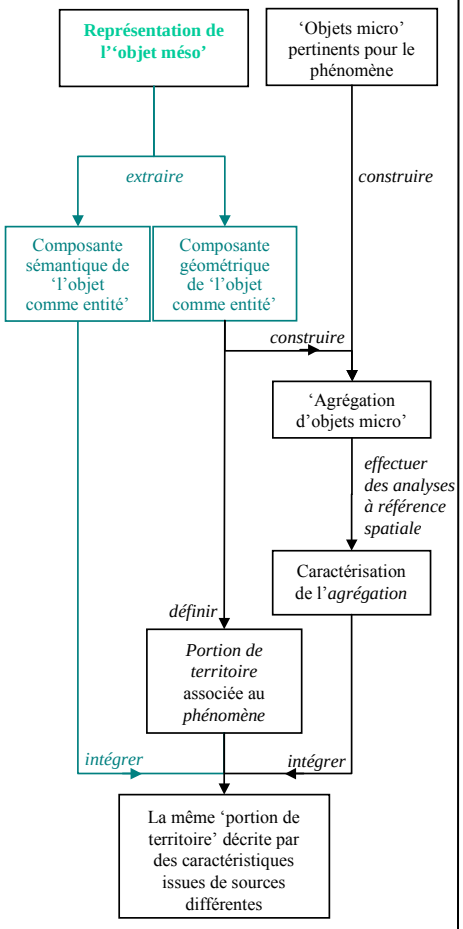
Figure B.2.36. *Schéma détaillé de la méthode d'observation multi-niveaux*

(double page suivante)

a- il existe une représentation de l'objet principal d'observation, sous forme d'information par nature : 'objet méso'

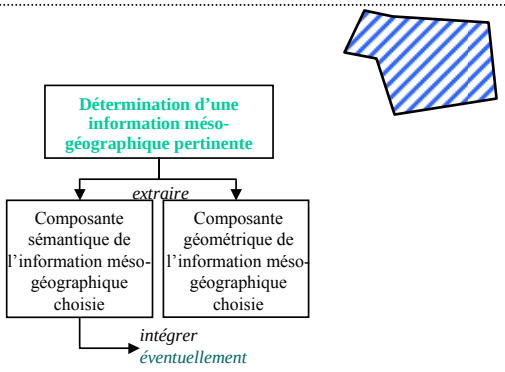
a- l'information méso-structurelle est prise égale à l''objet méso'

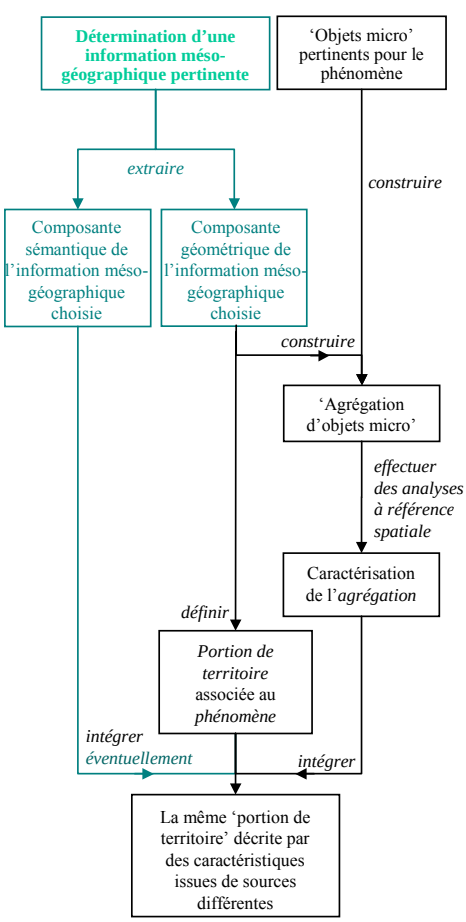




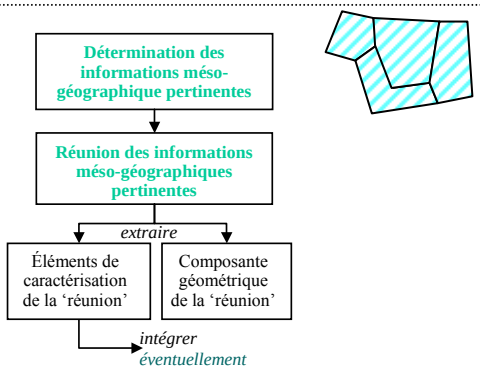
b- il n'existe pas de représentation de l''objet méso', MAIS il existe des informations méso-géographiques qui peuvent servir d'informations méso-structurelles

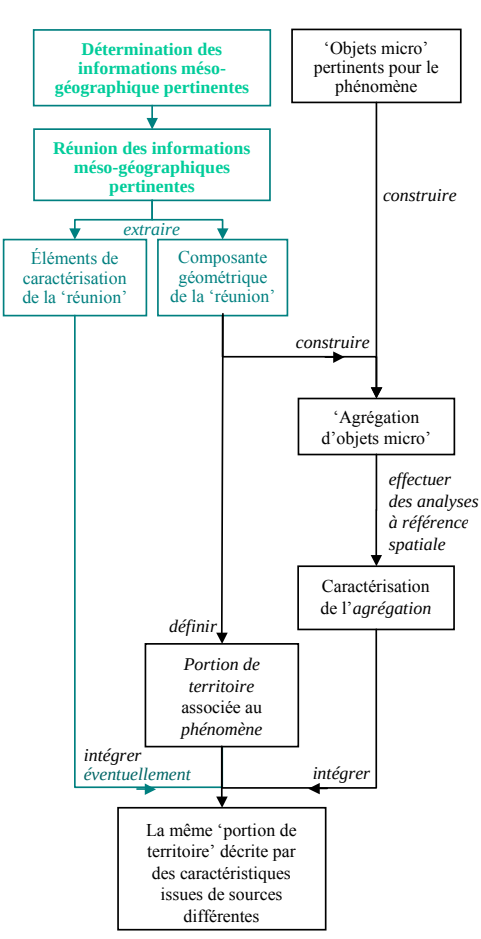
b1- l'information méso-structurelle est constituée d'*une seule* information méso-géographique


Détermination d'une information méso-géographique pertinente
extraire
Composante sémantique de l'information méso-géographique choisie
Composante géométrique de l'information méso-géographique choisie
intégrer éventuellement


Détermination d'une information méso-géographique pertinente
'Objets micro' pertinents pour le phénomène
extraire
construire
Composante sémantique de l'information méso-géographique choisie
Composante géométrique de l'information méso-géographique choisie
construire
'Agrégation d'objets micro'
effectuer des analyses à référence spatiale
Caractérisation de l'agrégation
définir
Portion de territoire associée au phénomène
intégrer éventuellement
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes

b2- l'information méso-structurelle est donnée sous la forme d'une *partition* d''objets méso-géographiques'

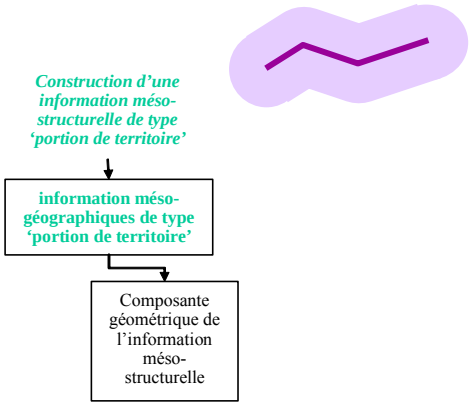

Détermination des informations méso-géographique pertinentes
Réunion des informations méso-géographiques pertinentes
extraire
Éléments de caractérisation de la 'réunion'
Composante géométrique de la 'réunion'
intégrer éventuellement


Détermination des informations méso-géographique pertinentes
'Objets micro' pertinents pour le phénomène
Réunion des informations méso-géographiques pertinentes
construire
extraire
Éléments de caractérisation de la 'réunion'
Composante géométrique de la 'réunion'
construire
'Agrégation d'objets micro'
effectuer des analyses à référence spatiale
Caractérisation de l'agrégation
définir
Portion de territoire associée au phénomène
intégrer éventuellement
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes

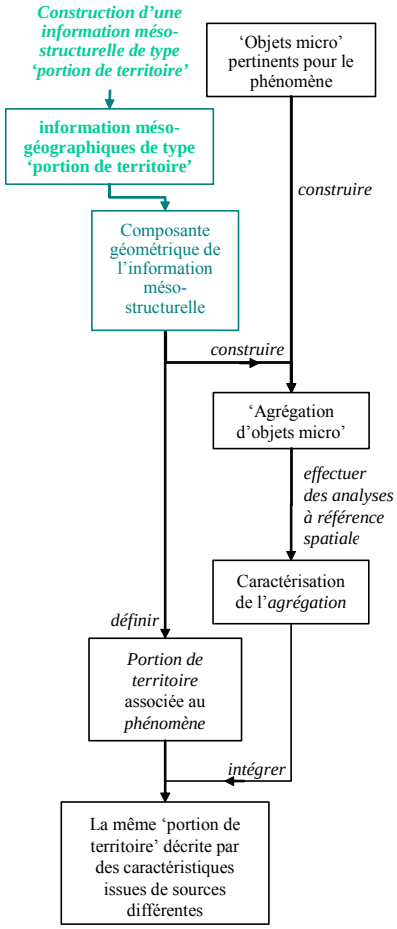
c- il n'existe pas d'information méso-géographique pour servir d'information méso-structurelle

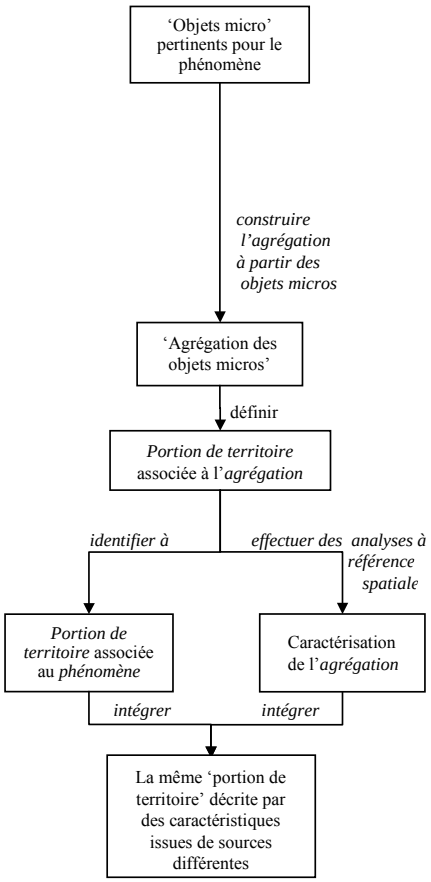
c1- Il existe une 'objet géographique' au niveau méso, à partir duquel on peut construire une information méso-structurelle de type 'portion de territoire' grâce à des fonctionnalités SIG (buffer), pour servir d' « emporte-pièce ».

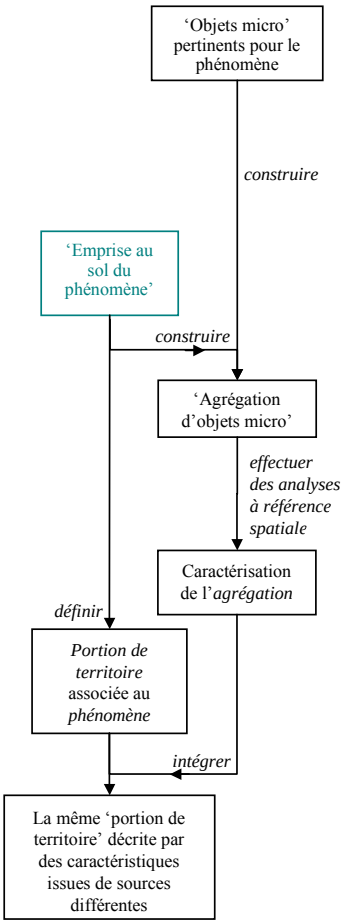
c2- où - on ne parvient pas à construire une information méso-géographique de type 'portion de territoire' pertinente, on utilise une agrégation à partir des 'objets micro'.


Construction d'une information méso-structurelle de type 'portion de territoire'
information méso-géographiques de type 'portion de territoire'
Composante géométrique de l'information méso-structurelle

Rappel : Application générale de la méthode à partir de l'emprise au sol


Construction d'une information méso-structurelle de type 'portion de territoire'
'Objets micro' pertinents pour le phénomène
information méso-géographiques de type 'portion de territoire'
construire
Composante géométrique de l'information méso-structurelle
construire
'Agrégation d'objets micro'
effectuer des analyses à référence spatiale
Caractérisation de l'agrégation
définir
Portion de territoire associée au phénomène
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes


'Objets micro' pertinents pour le phénomène
construire l'agrégation à partir des objets micros
'Agrégation des objets micros'
définir
Portion de territoire associée à l'agrégation
identifier à
effectuer des analyses à référence spatiale
Portion de territoire associée au phénomène
Caractérisation de l'agrégation
intégrer
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes


'Objets micro' pertinents pour le phénomène
construire
'Emprise au sol du phénomène'
construire
'Agrégation d'objets micro'
effectuer des analyses à référence spatiale
Caractérisation de l'agrégation
définir
Portion de territoire associée au phénomène
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes

B.2.5.3. Caractérisation des 'agrégations d'objets micro'

Après avoir montré comment construire des 'agrégations d'objets micro', nous allons maintenant nous intéresser à l'autre étape principale de la méthode d'observation multi-niveaux : la *caractérisation* des 'agrégations'. Les schémas des figures B.2.37 et B.2.38 montrent deux approches complémentaires de cette étape :

- l'une est globale et méthodologique. Elle décrit le processus de mise en œuvre de la méthode d'observation multi-niveaux et situe la caractérisation dans ce processus (fig. B.2.37),

- l'autre est technique et s'intéresse spécifiquement à la caractérisation dont elle décrit le principe en tant qu'opérations sur des données (fig. B.2.38).

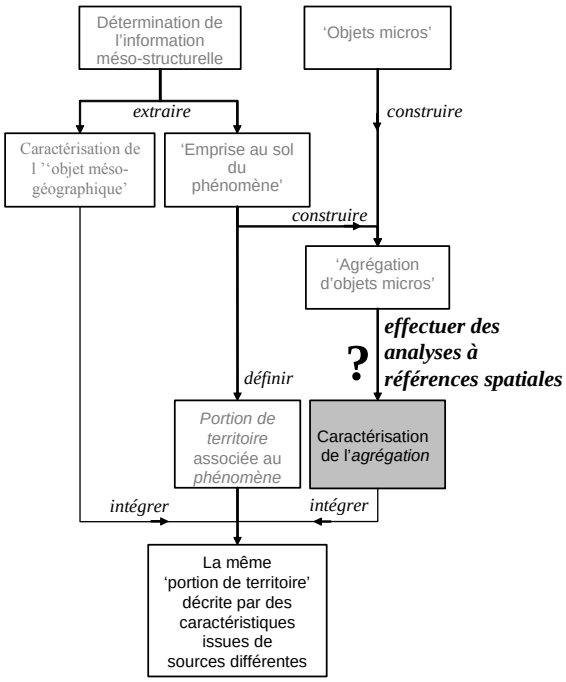
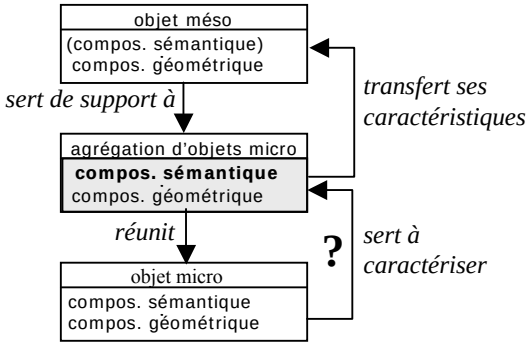


Figure B.2.37. *La caractérisation comme étape du processus d'observation multi-niveaux*



Comment caractériser les 'agrégations d'objets micro' pour transférer ensuite les indicateurs obtenus à l''objet méso géographique' (lorsqu'il existe) ou sinon à l'objet 'portion de territoire associée au phénomène'?

Figure B.2.38. *La caractérisation comme opérations sur les données gérées dans le SIG*

La *caractérisation* est une étape essentielle dans la mise en œuvre d'une observation multi-niveaux : elle élabore les informations qui serviront au suivi du phénomène étudié. Partant des 'agrégations d'objets micro' produites à l'étape précédente, la caractérisation effectue des analyses à référence spatiale sur les données micros afin de construire des informations pour caractériser ces 'agrégations'. Ces informations conçues comme des indicateurs sont ensuite associées directement au phénomène étudié pour enrichir son suivi, par le biais de la *portion de territoire* qu'il occupe. Dans le chapitre B.1, nous avons proposée une définition large des analyses à référence spatiale. L'enjeu de cette définition se trouve ici, dans la capacité à produire un grand nombre d'indicateurs variés.

B.2.5.3.1. Différentes caractérisations pour l'agrégation

Nous nous sommes inspirés des statistiques descriptives pour proposer deux premiers types de caractéristiques. Les premiers indicateurs que nous considérons servent à « résumer l'agrégation » et à caractériser « l'élément moyen » de l'agrégation. Ces indicateurs de tendances regroupent l'effectif de l'agrégation, le min et le max des valeurs quantitatives et les indicateurs dits de tendances centrales : la moyenne, la médiane, le mode.

Cette caractérisation s'inscrit pour nous dans la lignée de la quantification de la description des objets de [Franc, 2001, p.284] ; elle vient compléter les attributs sur l'*objet méso* dont pourrait disposer l'utilisateur par ailleurs. *« Cette quantification revient à associer à chaque objet étudié une liste de variables,… qui en simplifie la description sans la dénaturer. Les exemples les plus banals sont la superficie d'un territoire, qui rend compte de son étendue mais ignore l'organisation spatiale des terroirs, ou plus naturellement tout indicateur global d'une activité humaine comme le PNB,.. ».* Il importe de relever ici qu' *« il n'existe pas une seule façon de quantifier un objet ».* La définition des caractéristiques agrégées reste foncièrement fonction de ce que l'on cherche à observer et de celui qui observe, et doit se faire en collaboration avec un thématicien (fig. B.2.39). Par exemple, le nombre d'employés pourra être présenté sous forme de classes (ex. : 1, 2 à 3, 4 à 5, 6 à 10, plus de 10) plutôt que comme une moyenne (ex. : 2,6 employés). Autre exemple : Si chaque information géographique possède une date de création correspondant à la date d'intégration dans la base de données, plutôt que de faire une moyenne qui ne correspondrait pas à une mise à jour réelle, on préférera travailler sur l'effectif de données associées à chaque date de mise à jour.

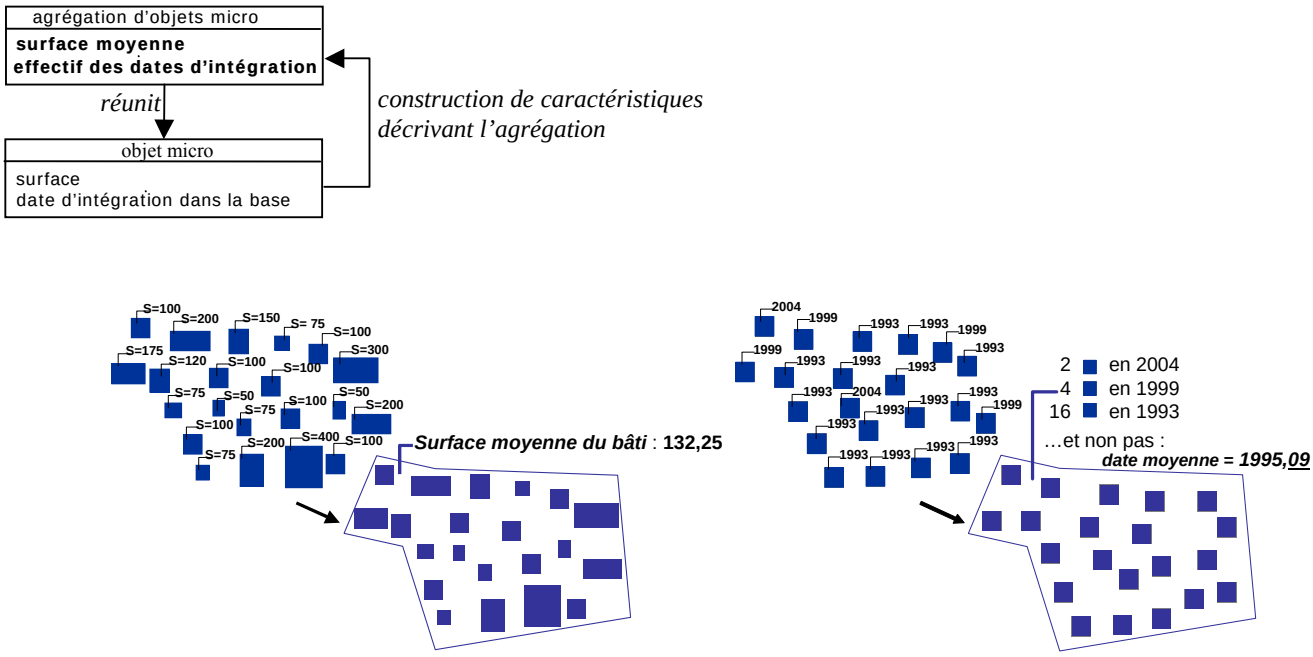


Figure B.2.39. *La forme prise par les indicateurs dépend de la construction des caractéristiques*

Construits à partir d'informations micro, ces indicateurs donnent des informations de tendance générale sur le niveau méso, qui est le niveau auquel on souhaite observer le phénomène. Toutefois, il est intéressant que *« les détails ne disparaissent pas au profit de formes globales et d'étendues plus larges, tandis que les niveaux les plus fins révèlent des différentiations spatiales locales »* [Mathian et Piron, 2001] citant [Durand-Dastès, 1999, p.64]. Nous proposons donc d'introduire des indicateurs d'homogénéité. La variance, par exemple, car *« l'analyse de la variance constitue la méthode la plus simple et la plus utilisée sans doute, … pour mesurer l'hétérogénéité interne d'un ensemble d'agrégat »* [Mathian et Piron, 2001, p.71]. Ce deuxième type de caractéristiques nous renseigne sur le caractère représentatif de « l'élément moyen ». Outre la variance ou l'écart type qui servent à caractériser la dispersion autour de la valeur centrale, nous utiliserons des indicateurs décrivant la forme de cette dispersion comme le

facteur d'aplatissement et de symétrie. Avec l'ensemble de ces indicateurs, nous seront alors à même de discriminer plus de situations. Par exemple, alors que chacune des zones présentées dans la figure B.2.40 comportent au total 20 entités élémentaires d'une surface moyenne égale à 100 m^2, elles traduisent des situations très différentes.

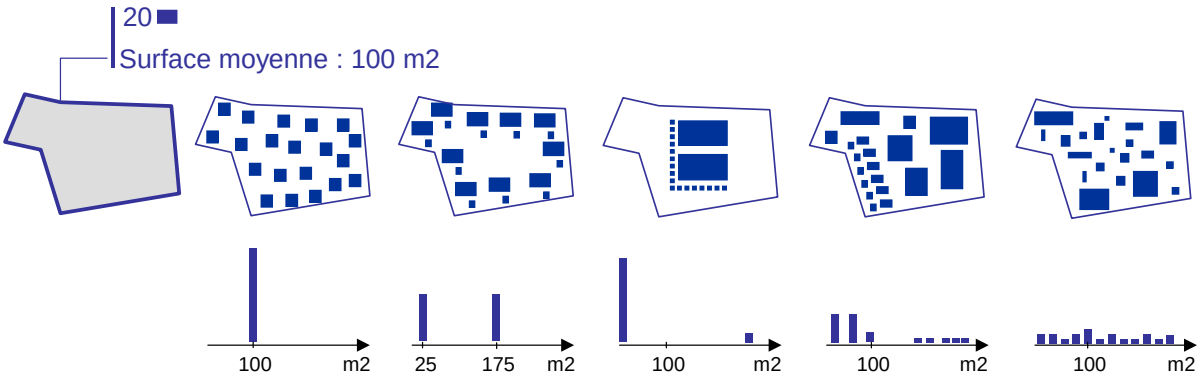


Figure B.2.40. *Une plus grande discrimination pour une observation plus détaillée*

Nous introduirons également un troisième groupe d'indicateurs plus spécifiques aux outils SIG qui exploitent principalement la composante géométrique des 'objets géographiques'. Ces indicateurs, que nous pourrions qualifier d'indicateurs « d'insertion territoriale », caractérisent la relation entre la surface de sol strictement occupée par l'agrégation et son «emprise» telle que nous l'avons définie. Comme nous construisons l''agrégation' à partir d'une 'emprise', cela a du sens d'étudier la relation entre cette 'emprise' et le territoire strictement occupé par l'ensemble des 'objets micro'. En revanche, cela n'en aurait pas eu autant si la stratégie d'agrégation adoptée était partie des 'objets micro', sans faire référence à une 'emprise méso' indépendante particulière. En effet quelle 'emprise' doit-on prendre dans ce cas comme support de référence pour effectuer les études d'« insertion » ? Enfin, si certains de ces indicateurs peuvent être calculés directement à partir d'informations sémantiques comme la densité (en effet la surface étant un attribut des 'objets géographiques', la densité peut être obtenue par simple ratio), nous envisageons ici surtout des analyses à partir de la composante géométrique comme par exemple l'analyse des espaces vides ou autres études sur la répartition des 'objets' dans l''emprise' (position, homogénéité, régularité, etc.). Ces indicateurs d'« insertion territoriale » viennent ainsi s'ajouter aux autres indicateurs pour permettre de distinguer encore plus de situations. La figure B.2.41 reprend l'exemple de la figure B.2.40 et montre comment les trois types d'indicateurs se complètent pour améliorer l'observation en proposant une discrimination plus fine des cas.

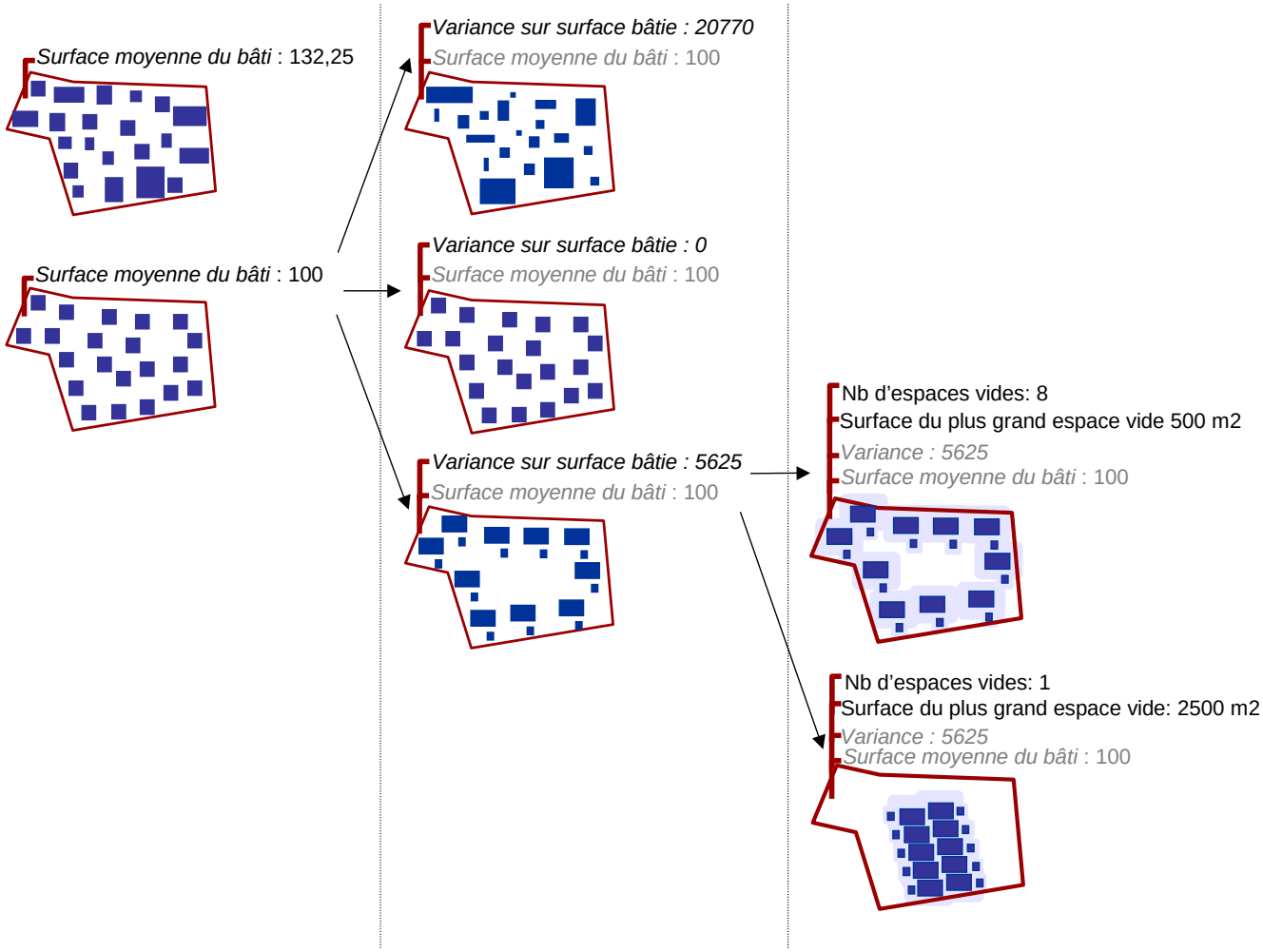

Surface moyenne du bâti : 132,25
Variance sur surface bâtie : 20770
Surface moyenne du bâti : 100
Surface moyenne du bâti : 100
Variance sur surface bâtie : 0
Surface moyenne du bâti : 100
Variance sur surface bâtie : 5625
Surface moyenne du bâti : 100
Nb d'espaces vides: 8
Surface du plus grand espace vide 500 m2
Variance : 5625
Surface moyenne du bâti : 100
Nb d'espaces vides: 1
Surface du plus grand espace vide: 2500 m2
Variance : 5625
Surface moyenne du bâti : 100

Figure B.2.41. *Mise en œuvre d'indicateurs variés pour différencier des situations proches*

B.2.5.3.2. Cas particulier où il existe une partition en informations méso-structurelles

En présentant les différentes formes d'information méso-structurelle, nous avons évoqué le cas (b2) où l'information méso-structurelle se présente sous la forme d'une partition d''objets méso'. Dans ce cas pour simplifier et nous ramener au cas général, nous suggérions de réunir les parties pour définir l'emporte-pièce nécessaire à la construction de l'agrégation. Nous proposions une mise en œuvre en trois étapes (cf. B.2.42) :

1. construction d'une représentation de l''emprise du phénomène' à partir de la réunion des 'emprises' des 'objets méso-structurels',
2. utilisation de cette 'emprise' (réunion) comme emporte-pièce pour agréger les 'objets micro',
3. caractérisation de l''agrégation d'objets micro' ainsi construite.

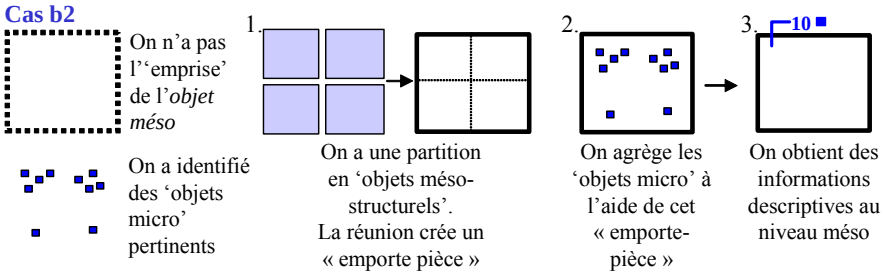
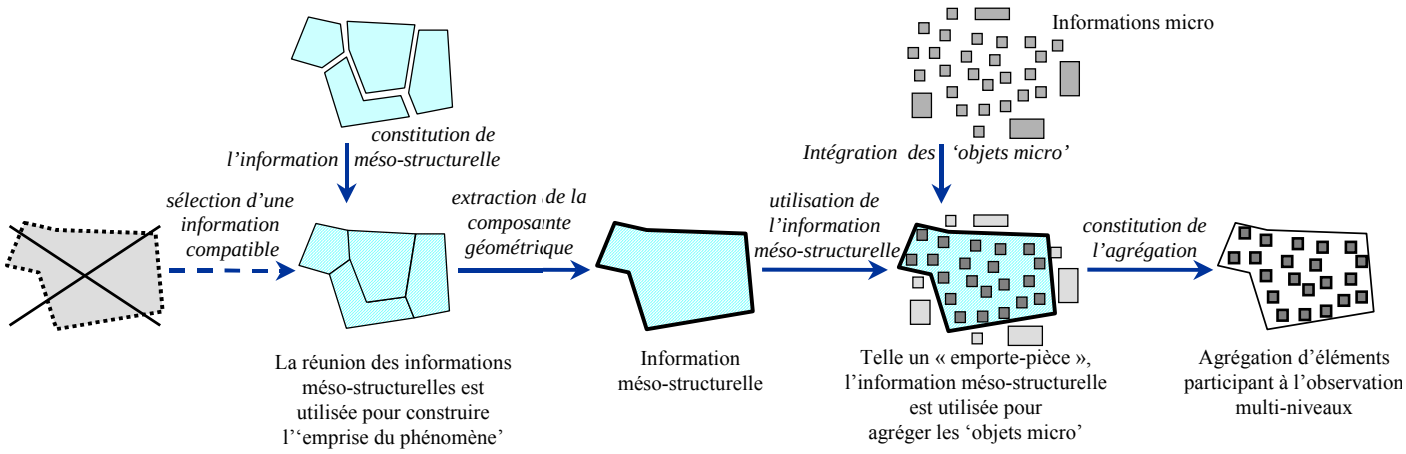

Cas b2
On n'a pas l''emprise' de l'objet méso
On a identifié des 'objets micro' pertinents
1.
On a une partition en 'objets méso-structurels'. La réunion crée un « emporte pièce »
2.
On agrège les 'objets micro' à l'aide de cet « emporte-pièce »
3.
10
On obtient des informations descriptives au niveau méso

Figure B.2.42. *Application basique de la méthode d'observation multi-niveaux dans le cas (b2) où il existe une information méso-structurelle sous forme de partition*



(rappel) *Principe général de la méthode d'observation dans le cas d'une information méso-structurelle sous forme de partition*

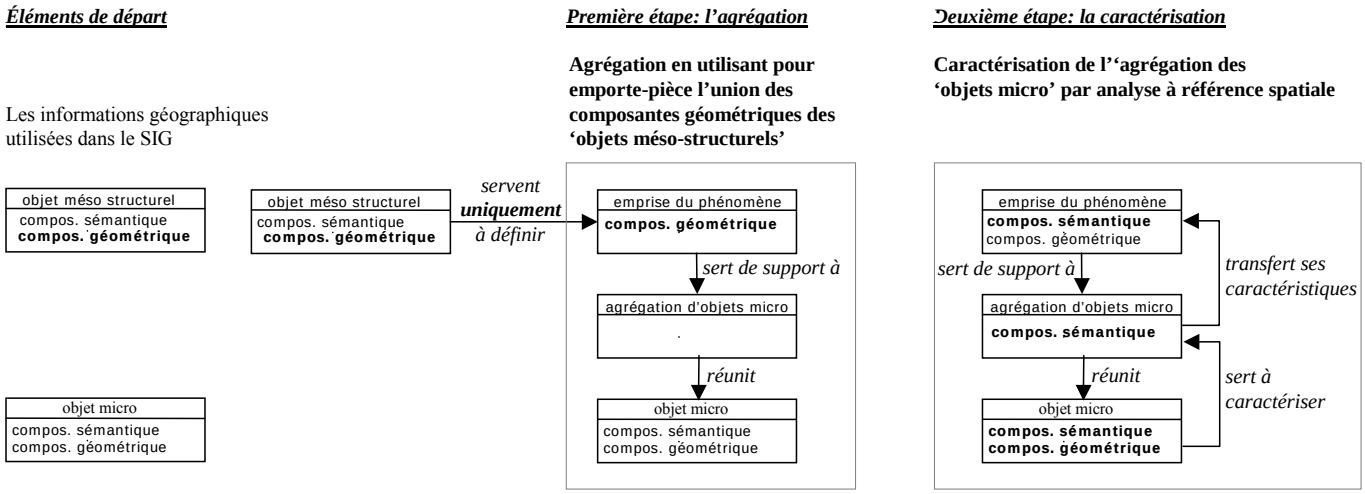


Figure B.2.43. *Traduction sur les données du principe général de la méthode d'observation multi-niveaux dans le cas où l''emprise' est donnée par l'union de composantes géométriques d'informations géographiques*

Dans le cas b2, on agrège les 'objets micro' sur la base d'une seule *emprise, qui correspond à l'emprise au sol du phénomène, donnée sous la forme d'une union de composantes géométriques d'informations géographiques dites 'méso-structurelles'.*

Au lieu de se ramener au cas général, il peut être plus judicieux d'exploiter autrement cette situation particulière ; nous proposons de ne pas considérer les informations méso-structurelles uniquement comme des *objets méthodologiques* et de ne pas les exploiter seulement pour définir l'emporte-pièce. Nous suggérons de les considérer comme des 'objets méso-géographiques' à part entière et de les observer comme tels. Nous conseillons en effet d'introduire un niveau d'observation supplémentaire (le niveau 2 venant s'ajouter au niveau 1) et un niveau d'agrégation additionnel (le niveau 2 s'ajoutent au niveau micro) (fig. B.2.44 et figure B.2.46.) et d'appliquer la méthode générale à la partition. Les 'objets méso-structurels' jouent alors individuellement le rôle d''objet méso' (cf. figure B.2.45 et ligne centrale de la figure B.2.47) :

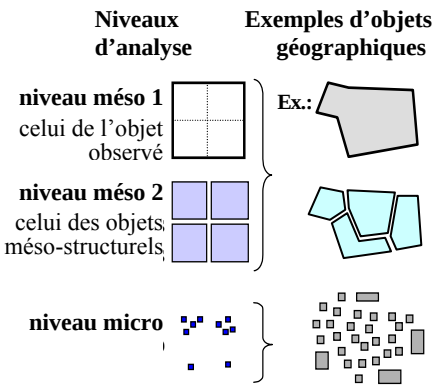


Figure B.2.44 . *Introduction du niveau méso 2*

Formellement, l'introduction du niveau méso des informations méso-structurelles, aboutit à trois niveaux d'observation différents. Cependant, dans le cas évoqué ici le niveau méso 1 n'existe pas (par hypothèse, cas b2).

1. sélection individuelle de la composante géométrique de chaque 'objet méso-structurel'
2. utilisation de ces 'emprises' pour agréger les 'objets micro',
3. caractérisation de chacune des 'agrégations' pour enrichir la caractérisation des 'objets méso-structurels'.

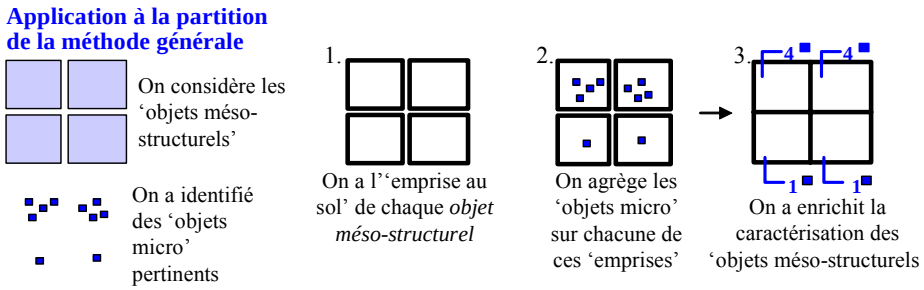

Application à la partition de la méthode générale
On considère les 'objets méso-structurels'
On a identifié des 'objets micro' pertinents
1.
On a l''emprise au sol' de chaque *objet méso-structurel*
2.
On agrège les 'objets micro' sur chacune de ces 'emprises'
3.
4
4
1
1
On a enrichit la caractérisation des 'objets méso-structurels

Figure B.2.45. *Première partie de la méthode d'observation multi-niveaux dans le cas où il existe une information méso-structurelle sous la forme d'une partition : application de la méthode générale à la partition*

Dans un second temps, les 'objets méso-structurels' jouent alors le rôle d'informations micro, et sont agrégés selon l''emprise du phénomène' (donnée par l'union de leur composante géométrique) (cf. figure B.2.47). Cette agrégation est à son tour caractérisée pour fournir des indicateurs au niveau d'observation principal, le niveau méso 1 (cf. figure B.2.48).

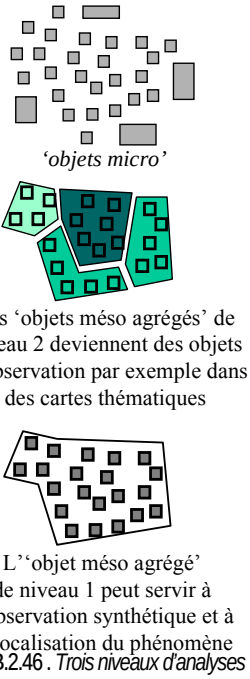

'objets micro'
Les 'objets méso agrégés' de niveau 2 deviennent des objets d'observation par exemple dans des cartes thématiques
L''objet méso agrégé' de niveau 1 peut servir à l'observation synthétique et à la localisation du phénomène

Figure B.2.46 . *Trois niveaux d'analyses*

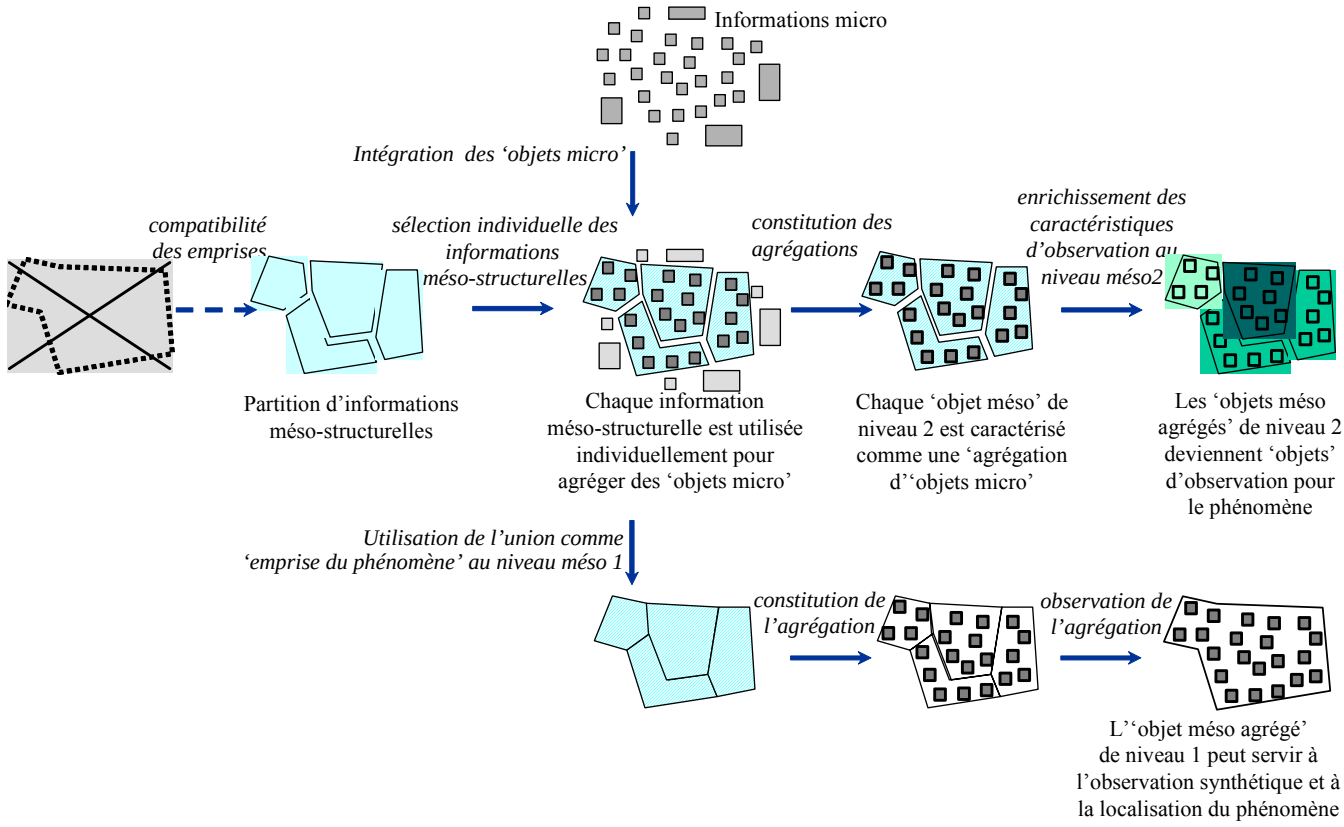

Informations micro
Intégration des 'objets micro'
compatibilité des emprises
sélection individuelle des informations méso-structurelles
constitution des agrégations
enrichissement des caractéristiques d'observation au niveau méso2
Partition d'informations méso-structurelles
Chaque information méso-structurelle est utilisée individuellement pour agréger des 'objets micro'
Chaque 'objet méso' de niveau 2 est caractérisé comme une 'agrégation d''objets micro'
Les 'objets méso agrégés' de niveau 2 deviennent 'objets' d'observation pour le phénomène
Utilisation de l'union comme 'emprise du phénomène' au niveau méso 1
constitution de l'agrégation
observation de l'agrégation
L''objet méso agrégé' de niveau 1 peut servir à l'observation synthétique et à la localisation du phénomène

Figure B.2.47. *Application de la méthode d'observation pour une observation à plus de deux niveaux*

Dans une première étape, l'application de la méthode d'observation multi-niveaux aux 'objets méso-structurels' permet d'enrichir leurs attributs de caractéristiques relatives au phénomène étudié et ainsi de pouvoir effectuer des analyses à ce niveau méso intermédiaire (par exemple sous forme de cartes thématiques, cf. ligne centrale de la figure). Ensuite, l'union des composantes géométriques de ces 'objets' permet de construire des 'agrégations d'objets micro' pour caractériser le niveau méso 1 qui reste le niveau d'observation principal : celui qui permet l'observation et le suivi direct du phénomène dans le SIG (cf. la ligne inférieure).

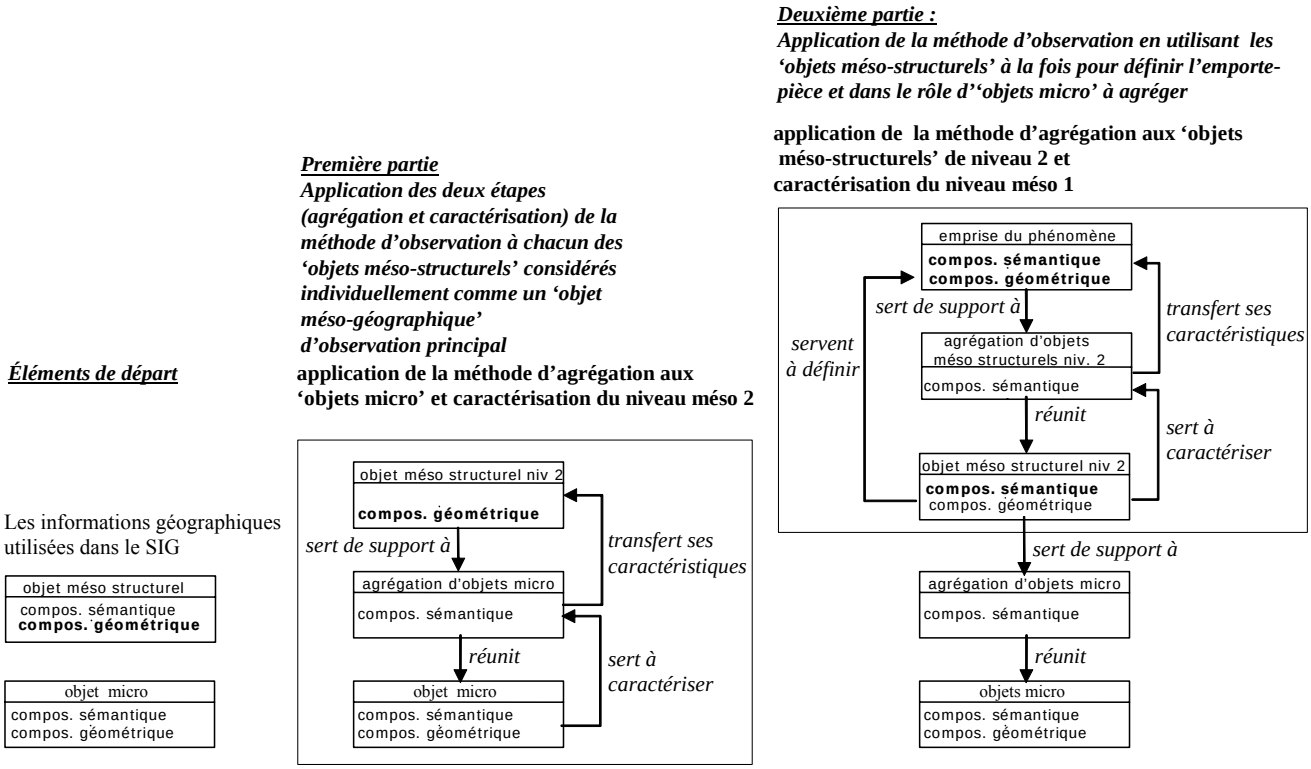

Deuxième partie :
Application de la méthode d'observation en utilisant les 'objets méso-structurels' à la fois pour définir l'emporte-pièce et dans le rôle d''objets micro' à agréger
application de la méthode d'agrégation aux 'objets méso-structurels' de niveau 2 et caractérisation du niveau méso 1
Première partie
Application des deux étapes (agrégation et caractérisation) de la méthode d'observation à chacun des 'objets méso-structurels' considérés individuellement comme un 'objet méso-géographique' d'observation principal
Éléments de départ
application de la méthode d'agrégation aux 'objets micro' et caractérisation du niveau méso 2
Les informations géographiques utilisées dans le SIG
objet méso structurel
compos. sémantique
compos. géométrique
objet micro
compos. sémantique
compos. géométrique
objet méso structurel niv 2
compos. géométrique
sert de support à
agrégation d'objets micro
compos. sémantique
réunit
objet micro
compos. sémantique
compos. géométrique
transfert ses caractéristiques
sert à caractériser
emprise du phénomène
compos. sémantique
compos. géométrique
sert de support à
servent à définir
agrégation d'objets méso structurels niv. 2
compos. sémantique
réunit
objet méso structurel niv 2
compos. sémantique
compos. géométrique
transfert ses caractéristiques
sert à caractériser
sert de support à
agrégation d'objets micro
compos. sémantique
réunit
objets micro
compos. sémantique
compos. géométrique

Figure B.2.48. *Schéma des opérations composant la méthode d'observation multi-niveaux dans le cas on l'on souhaite effectuer une analyse à plusieurs niveaux*

Cette mise en œuvre différenciée de la méthode d'observation multi-niveaux pour le cas où l'information méso-structurelle est donnée sous forme d'une partition permet :

- d'obtenir les mêmes indicateurs au niveau méso 1 que l'application standard

- d'élaborer des indicateurs supplémentaires issus de la caractérisation des agrégations du niveau 2. En particulier, elle offre la possibilité de conserver l'information d'une homogénéité des parties au sein d'un tout qui ne le serait pas. Par exemple si une 'zone pavillonnaire' réunit plusieurs 'îlots' on peut étudier de façon indépendante les groupes de bâtiment situés dans chacun de ces 'îlots' (voir figure B.2.49). Cela peut nous permettre *«d'extraire les structures que l'on perd dans le passage d'un niveau à l'autre»* [Mathian, Piron, 2001, p.72].

De plus, cela nous permettra par la suite d'observer de façon différenciée l'évolution de chaque partie et d'étudier leur participation à l'évolution globale du phénomène (par exemple est-ce que l'ensemble est homogène ou bien chaque partie a t'elle une description spécifique ?) pour répondre ensuite lors du suivi à des questions du type : Une seule partie porte t'elle seule les évolutions du tout ?

- Enfin, le caractère « emboîté » des différents niveaux nous place dans un contexte particulier. Des travaux ultérieurs pourront chercher à mieux l'exploiter. En particulier, il s'agira de trouver des indicateurs pour mettre en évidence des différenciations entre les sous-groupes d''objets micro' définis par les 'agrégations' de niveau méso 2 au sein de l''agrégation' de niveau 1 pour tenter de saisir les éventuelles structures induites par les 'objets méso-structurels' que nous ne distinguons pas sinon.

Figure B.2.49. *Etude de 'îlots' participant à une* 'zone pavillonnaire'

Expérience issue de l'étude sur les zones pavillonnaires (cf. Partie C).

Une mise en œuvre différentiée de la méthode visant à identifier et à caractériser les 'zones pavillonnaires' dans la BD TOPO® à partir des 'îlots routiers' a été testée dans le stage d'O. Raimond que nous avons encadré. Son objectif secondaire consistait à établir une typologie des 'zones pavillonnaires' proche de catégories architecturales à l'aide d'indicateurs morphologiques. Cette carte extraite de [Raimond, 2005] est une illustration des premiers résultats obtenus.

B.2.6. Extensions de la méthode d'observation multi-niveaux

Partant du besoin d'informations pour observer un phénomène, nous avons élaboré une méthode pour en produire. Elle propose dans un premier temps une solution au problème technique de mise en relation de représentations différentes en informations géographiques ; ce qui, dans un second temps, lui permet d'exploiter la complémentarité de ces représentations pour construire de nouveaux indicateurs à observer (fig. B.2.50).

La méthode qui utilise une information méso-structurelle pour agréger les 'objets micro' prend différentes formes en fonction du contexte de l'utilisateur. Cette différence de forme ne se traduit pas par une gradation des difficultés techniques de mise en œuvre; celle-ci restant toujours la même et se réduisant essentiellement à une requête géométrique basique de type : « est inclus » ②. En revanche, elle impose à l'utilisateur un travail d'analyse plus important pour déterminer des informations méso-structurelles pertinentes pour le phénomène①.

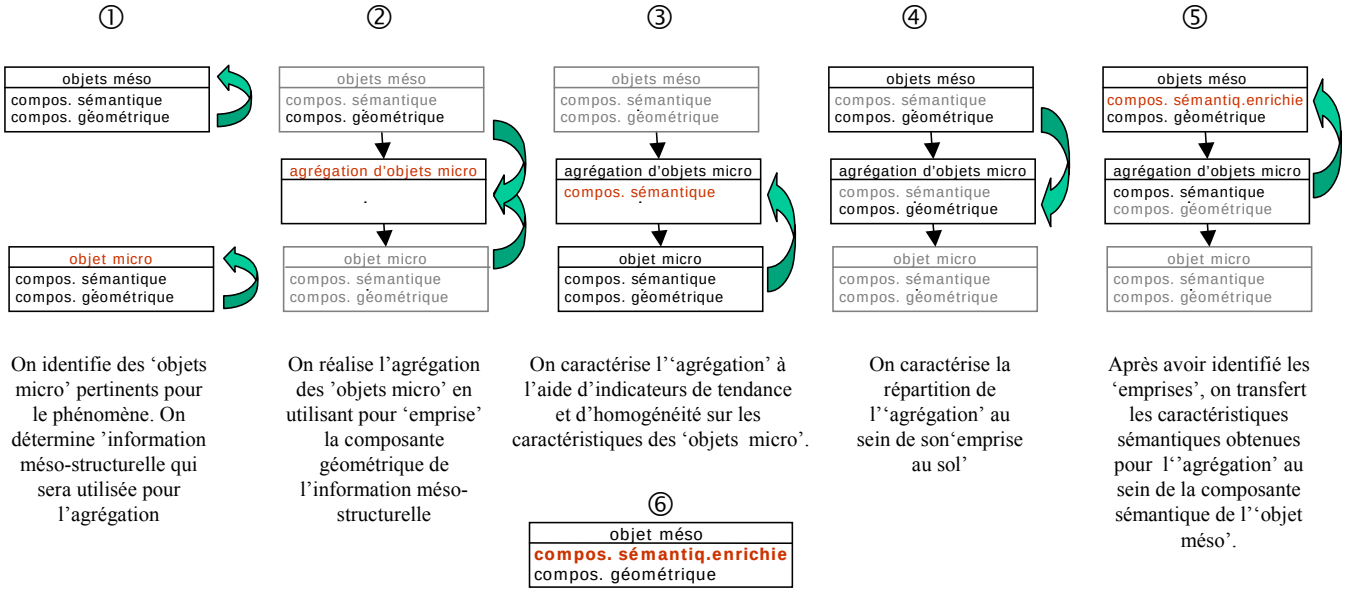

①
②
③
④
⑤
objets méso
compos. sémantique
compos. géométrique
agrégation d'objets micro
objet micro
compos. sémantique
compos. géométrique
compos. sémantiq.enrichie
On identifie des 'objets micro' pertinents pour le phénomène. On détermine 'information méso-structurelle qui sera utilisée pour l'agrégation
On réalise l'agrégation des 'objets micro' en utilisant pour 'emprise' la composante géométrique de l'information méso-structurelle
On caractérise l''agrégation' à l'aide d'indicateurs de tendance et d'homogénéité sur les caractéristiques des 'objets micro'.
On caractérise la répartition de l''agrégation' au sein de son'emprise au sol'
Après avoir identifié les 'emprises', on transfert les caractéristiques sémantiques obtenues pour l''agrégation' au sein de la composante sémantique de l''objet méso'.
⑥
objet méso
compos. sémantiq.enrichie
compos. géométrique
Au final, on obtient bien un enrichissement de la représentation l''objet méso' servant à observer le phénomène étudié. Lorsque cette représentation n'existe pas , la méthode permet d'enrichir la description de la ' portion de territoire' concernée par le phénomène pour tenter de l'observer au travers de ses effets sur cette portion de territoire.

Figure B.2.50. *Présentation des différentes étapes de la méthode sous forme d'opérations sur des données dans un SIG*

Nous avons veillé à ce que chacune des étapes du processus d'observation puisse être effectuée avec des fonctionnalités standards d'un SIG. Cela lui assure d'être simple à mettre en œuvre, mais cette simplicité n'implique pas pour autant de renoncer au potentiel d'analyse des SIG. La méthode peut également mettre à profit des compétences, outils et données supplémentaires dont bénéficierait l'utilisateur par ailleurs. En effet, nous visions l'élaboration d'une méthode à la fois ouverte (pour tirer parti de l'existant et satisfaire des utilisateurs aux profils variés) et évolutive (pour accompagner les utilisateurs dans le développement de leurs observations). En la concevant ouverte sur l'existant (existant en données, mais aussi en modèles de mise à jour plus développés, en méthodes et outils d'analyses à référence spatiale plus spécialisés), la méthode devient de fait *adaptative* ; il lui suffit pour cela d'intégrer des moyens et compétences

supplémentaires. Ainsi, la méthode est facile à mettre en œuvre, mais n'est pas limitée par (ou à) l'exploitation de requêtes élémentaires. La méthode d'observation peut tirer parti des moyens disponibles pour enrichir la description des phénomènes étudiés. En contre-partie toutefois, elle implique souvent des investissements complémentaires (sous forme de données, de développements ou d'acquisition de compétences en analyses à référence spatiale). L'utilisateur suivant son profil, son contexte et ses ressources privilégiera certains modes d'enrichissement plutôt que d'autres ; tous cependant ne possèdent pas le même potentiel. Au final, le produit de la caractérisation dépendra beaucoup de l'utilisateur et des moyens dont il dispose (pour calculer des caractéristiques additionnelles ou intégrer des données supplémentaires).

- enrichissement de la description de l''objet méso' par intégration dans la composante sémantique de données supplémentaires (cf. B.2.6.1),
- exploitation d'analyses morphologiques (cf. B.2.6.2),
- exploitation d'analyses topologiques (cf. B.2.6.3),
- exploitation d'analyses statistiques et prise en compte d'un niveau d'observation intermédiaire (cf. B.2.6.4),
- construction de plusieurs 'agrégations d'objets micro' (cf. B.2.6.5).

B.2.6.1. Enrichissement de la composante sémantique des 'objets' utilisés

Une façon simple d'enrichir la description du phénomène observé consiste à intégrer de nouveaux attributs à la composante sémantique des 'objets géographiques' étudiés. Cet enrichissement peut s'appliquer à :

- à l'information géographique qui sert à définir l'«emporte-pièce» ('objet-méso' ou autre 'information méso-structurelle'). Cette opération vient alors compléter directement la caractérisation de l''emprise' du phénomène (cf. figure B.2.51),
- aux 'objets micro'. Cette opération permet d'enrichir la liste des indicateurs produits pour caractériser l'*agrégation* et donc indirectement le phénomène étudié (cf. figure B.2.52).

Par exemple, pour étudier le développement d'une *zone industrielle* à partir de la BD TOPO®, on peut ajouter à l''enceinte industrielle' telle qu'elle est décrite dans la base : une 'date de création', un 'statut juridique', des 'contraintes d'urbanisme'… Et parallèlement, on peut introduire des informations sur l'activité de chacune des 'entreprises' situées dans cette 'zone' (ex. : chiffre d'affaire, nombre d'employés, secteur d'activité).

Cet enrichissement s'effectue sous la forme d'une « jonction de fichier » à l'aide d'un identifiant (ou clé primaire, ou index) et plus généralement par géocodage. Cette opération ne pose pas de grandes difficultés techniques. Son coût principal réside dans l'acquisition de données supplémentaires.

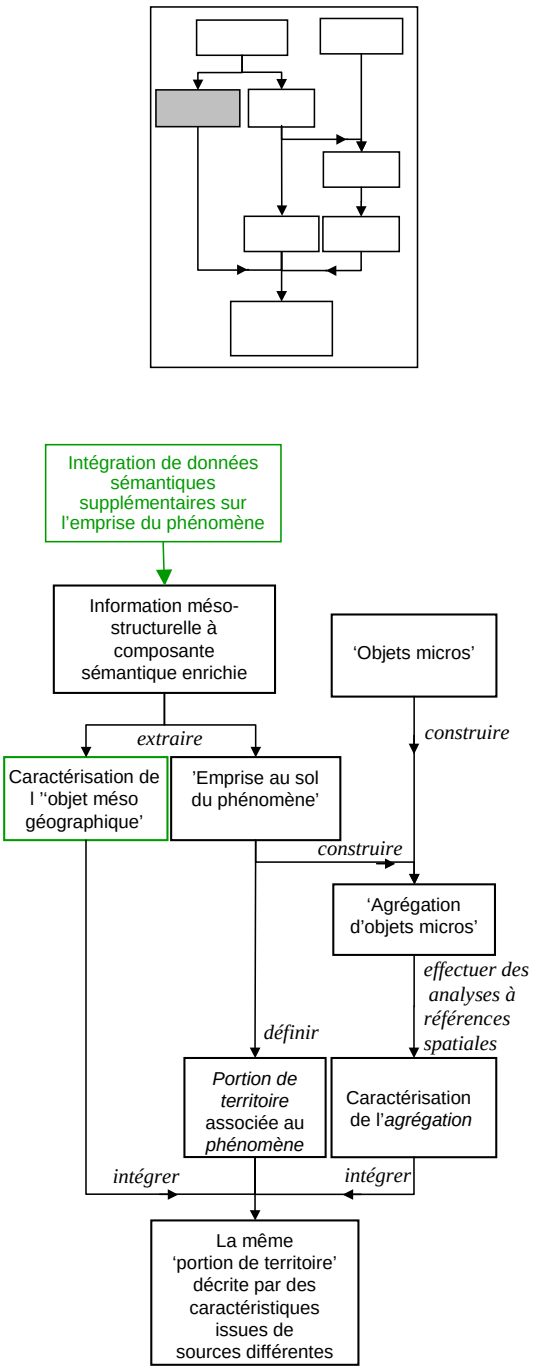

Intégration de données sémantiques supplémentaires sur l'emprise du phénomène
Information méso-structurelle à composante sémantique enrichie
'Objets micros'
extraire
construire
Caractérisation de l "objet méso géographique'
'Emprise au sol du phénomène'
construire
'Agrégation d'objets micros'
effectuer des analyses à références spatiales
définir
Portion de territoire associée au phénomène
Caractérisation de l'agrégation
intégrer
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes

Figure B.2.51. *Enrichissement de la méthode par ajout d'informations descriptives sur l''objet méso'*

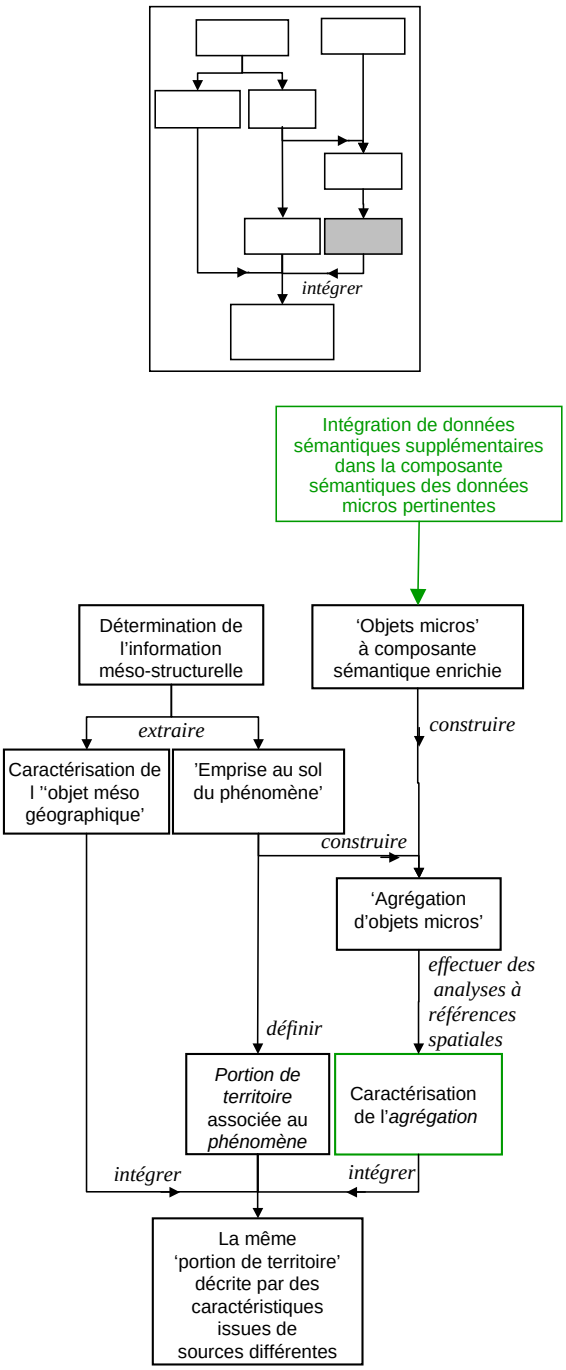

intégrer
Intégration de données sémantiques supplémentaires dans la composante sémantiques des données micros pertinentes
Détermination de l'information méso-structurelle
'Objets micros' à composante sémantique enrichie
extraire
construire
Caractérisation de l "objet méso géographique'
'Emprise au sol du phénomène'
construire
'Agrégation d'objets micros'
effectuer des analyses à références spatiales
définir
Portion de territoire associée au phénomène
Caractérisation de l'agrégation
intégrer
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes

Figure B.2.52. *Enrichissement de la méthode par ajout d'informations descriptives sur les 'objets micro'*

B.2.6.2. Exploitation d'analyses morphologiques

La composante géométrique contient implicitement de nombreuses informations sur la forme des *objets géographiques* (fig.B.2.53). Les analyses morphologiques permettent de les extraire ([Regnault et al, 1999], [Ruas, 1999], [Duchène, 2004]).

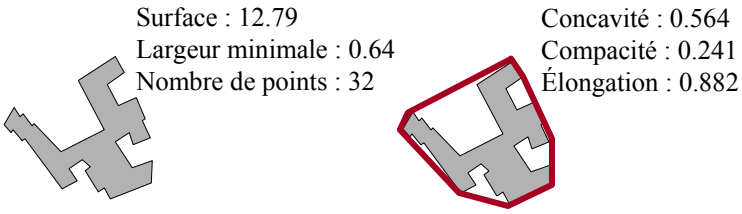


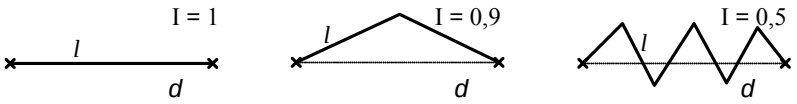
Figure B.2.53. *Description d'un polygone (bâtiment) [Regnauld et al, 1999]*

Leur explicitation nécessite généralement des développements De fait, les analyses morphologiques restent souvent sous-employées et le potentiel de la composante géométrique sous-exploité par les utilisateurs de SIG. Pourtant, il existe des analyses simples à mettre en œuvre. Par exemple (fig. B.2.54 et fig.B.2.55).

On peut quantifier la « rectilignité » d'un tronçon de route en utilisant le ratio :

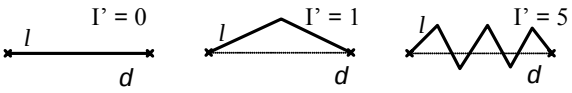
$$\frac{\text{Distance entre les deux points extrémités du tronçon}}{\text{longueur totale du tronçon donnée par le SIG}}$$

Indicateur de « rectilignité » $I = d / l$

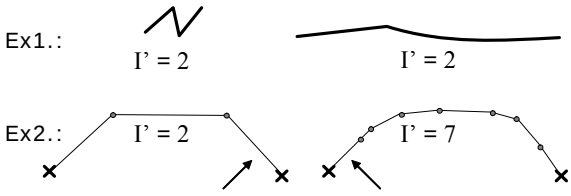


En revanche, le nombre de points intermédiaires dans le tracé de la ligne qui pourrait sembler un bon indicateur, ne convient pas ; il dépend de la longueur de la polyligne et même en introduisant une pondération par cette longueur, la pratique montre qu'un même objet selon l'opérateur qui le numérise peut avoir une résolution spatiale assez variable.

Le nombre de points intermédiaires…



…comme mauvais indicateur de « rectilignité ».



même tronçon numérisé par deux personnes différentes

Figure B.2.54. *Un bon et un mauvais indicateur de la « rectilignité »*

Indicateur de complexité des formes:

C= (périmètre)² /surface

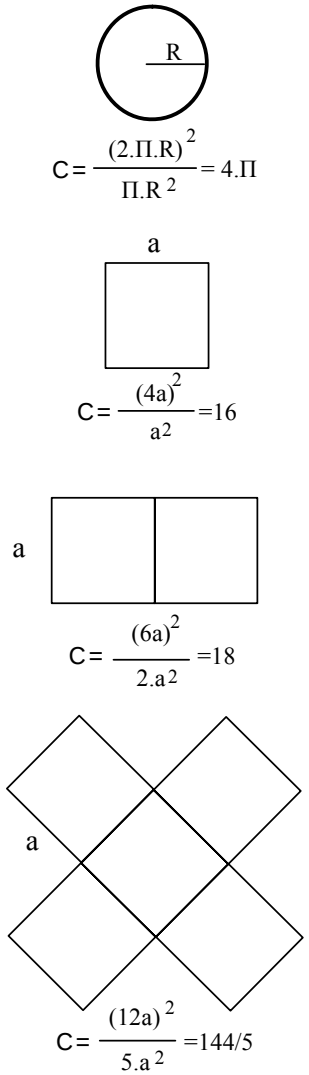


Figure B.2. 55. *Exemples d'indicateurs de complexité de forme*

Si ces premiers indicateurs sont insuffisants pour répondre aux besoins, l'utilisateur pourra alors se tourner vers des indicateurs plus élaborés, qu'il trouvera présentés dans les travaux de recherches dédiées (par exemple les travaux en généralisation de [Regnault, 1998], ceux de [Duchêne et al., 2003] sur l'orientation des bâtiments, [Fritsch, 1997] sur la sinuosité des tronçons de routes, etc.).

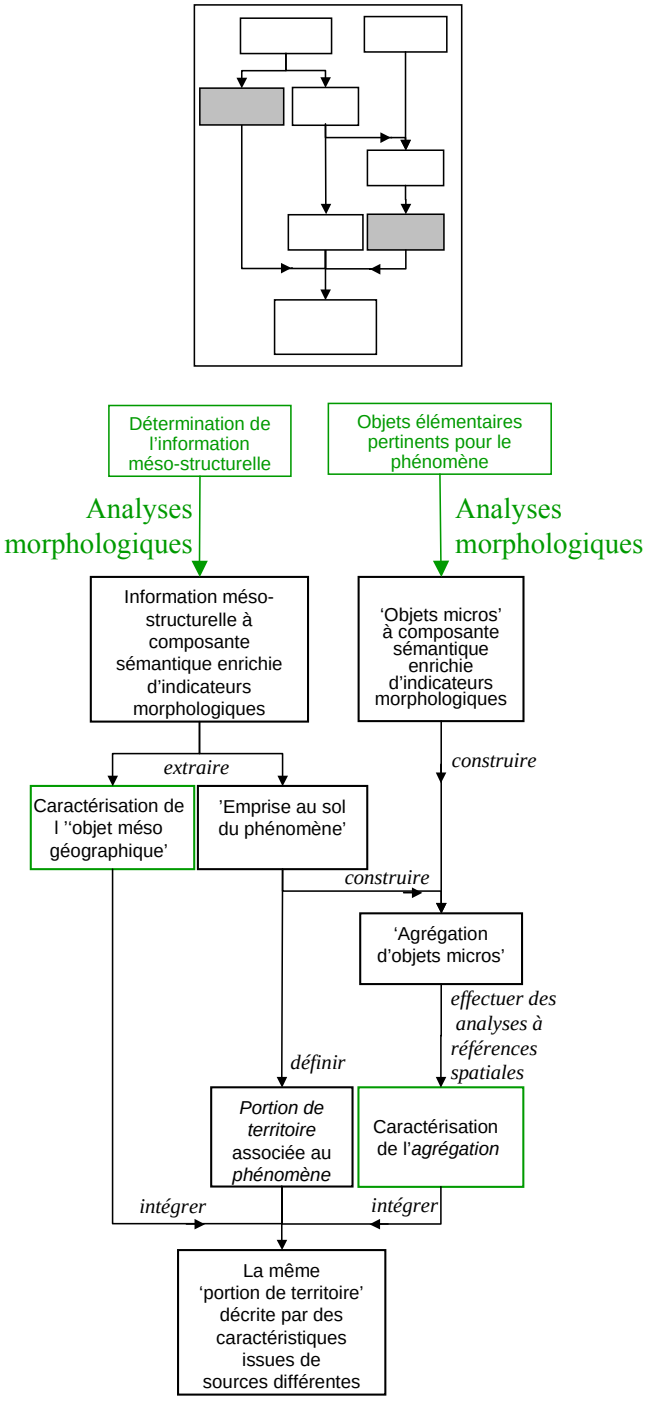

Détermination de l'information méso-structurelle
Objets élémentaires pertinents pour le phénomène
Analyses morphologiques
Analyses morphologiques
Information méso-structurelle à composante sémantique enrichie d'indicateurs morphologiques
'Objets micros' à composante sémantique enrichie d'indicateurs morphologiques
extraire
construire
Caractérisation de l ''objet méso géographique'
'Emprise au sol du phénomène'
construire
'Agrégation d'objets micros'
effectuer des analyses à références spatiales
définir
Portion de territoire associée au phénomène
Caractérisation de l'agrégation
intégrer
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes

Figure B.2.56. *Enrichissement de la méthode grâce à l'utilisation d'analyses morphologiques pour enrichir la description des 'objets' (méso et/ou micro)*

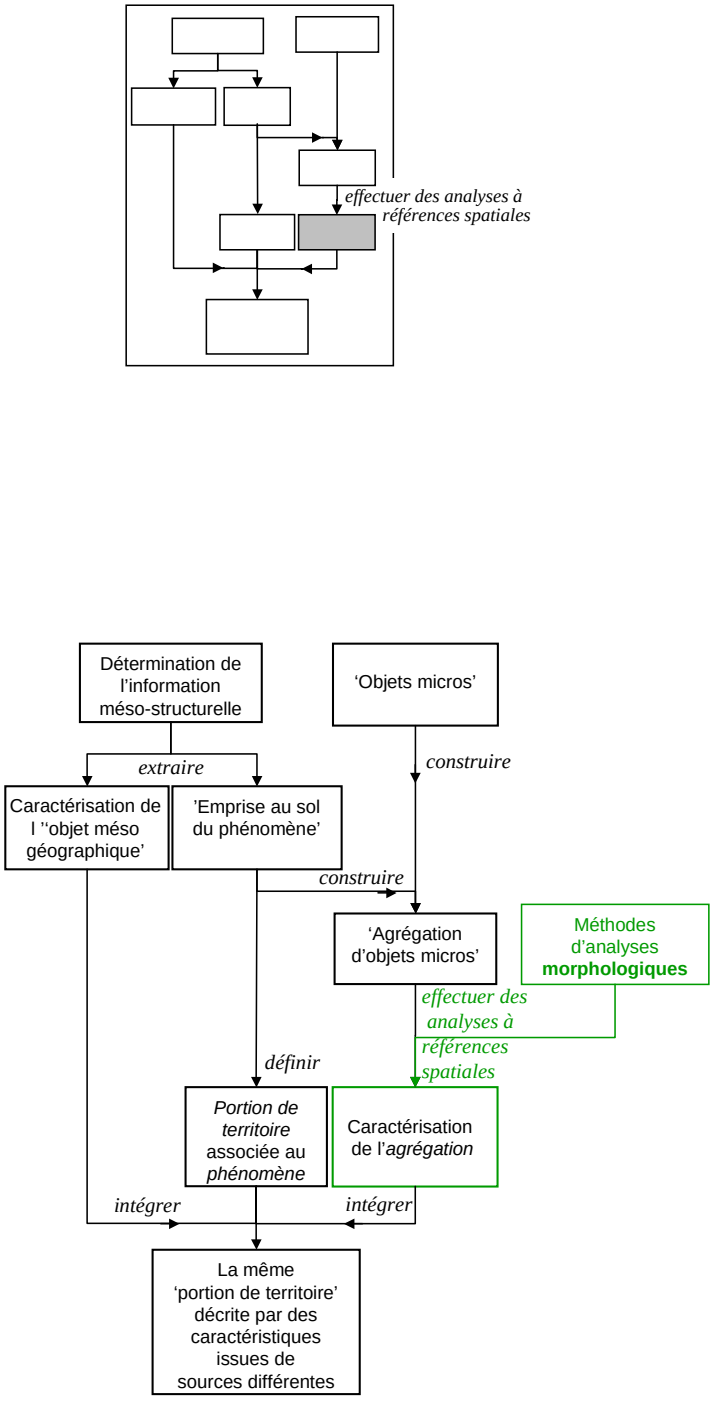


Figure B.2.57. *Enrichissement de la méthode par intégration d'analyses morphologiques dans la caractérisation des 'agrégations'*

La valeur de ces indicateurs, une fois calculée, sera intégrée à la composante sémantique de chaque objet. Si les objets concernés sont des 'objets' du niveau méso (fig. B.2.56), la caractérisation supplémentaire issue de l'analyse morphologique vient directement enrichir la description

du phénomène. Si ce sont des 'objets micro' (fig. B.2.57), alors par agrégation ces indicateurs serviront à construire des informations de tendance (ex. : élongation moyenne, concavité moyenne, « rectilignité » moyenne) et des informations d'homogénéité.

Introduire des analyses morphologiques en complémentarité avec l'application de la méthode d'observation multi-niveaux revient à travailler avec des données initiales plus riches. Exploiter ces analyses dans la phase de caractérisation de l'agrégat permet d'obtenir par ailleurs une gamme plus large d'indicateurs (cf. figure B.2.57), en particulier des indicateurs sur la répartition de l'agrégation dans son emprise. En effet, ces indicateurs étant par définition essentiellement géométriques, les analyses morphologiques sont les moyens spécifiques pour les construire (ex.: la qualification des espaces « vides », cf. fig. B.2.58).

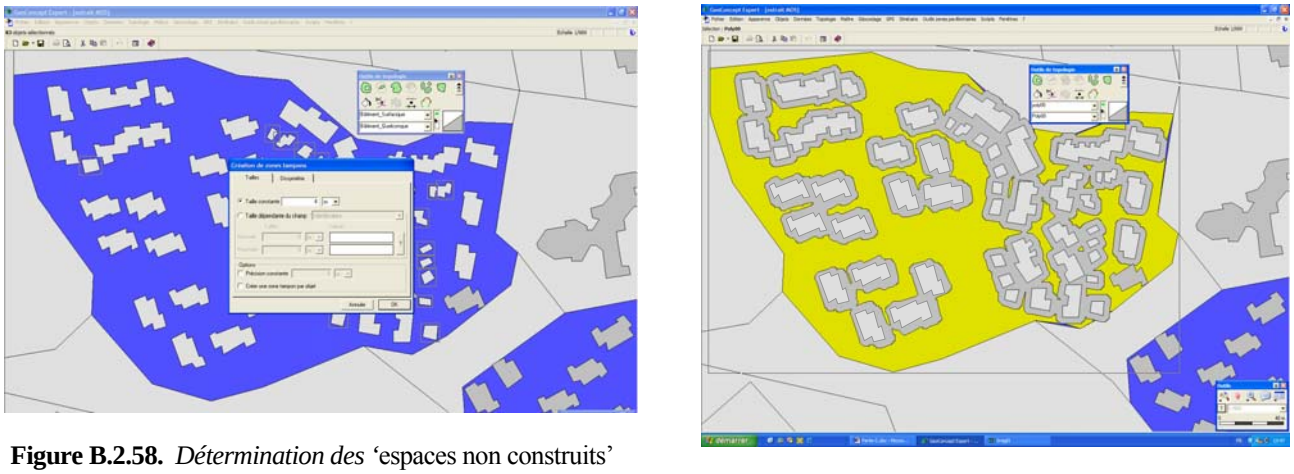


Figure B.2.58. *Détermination des* 'espaces non construits'

1. Sélection de tous les bâtiments à l'intérieur de la zone
2. Construction d'un buffer autour de chacun des bâtiments (valeur prise ici à 10, cette valeur fait partie des paramètres qu'il faut déterminer avec l'aide d'un thématicien)
3. Construction de 'la surface construite' *par agrégation des buffers particuliers à chaque bâtiment*
4. Construction par complémentarité à l'intérieur de la zone, de 'la surface non

La difficulté de ces enrichissements réside dans l'élaboration et la mise en œuvre des indicateurs morphologiques. Cette difficulté peut être diminuée en mettant à profit les fonctionnalités existantes et les résultats déjà obtenus en recherche. Le coût principal sera alors celui induit par les développements.

B.2.6.3. Exploitation d'analyses topologiques

La composante topologique, comme la composante géométrique, est souvent sous-exploitée, malgré la richesse d'informations qu'elle contient. Des travaux comme ceux de [Gleyze, 2005] (sur la « détermination des points de vulnérabilité d'un réseau ») ou de [Claramunt, 2005] (sur la « représentation d'une certaine logique de la cité ») montrent que des informations sophistiquées sur les relations spatiales des objets peuvent être obtenues. Cependant, il existe peu de fonctionnalités standards dans les SIG pour réaliser ce type d'étude (ex. : des fonctionnalités de calcul d'itinéraire sont proposées par les constructeurs, mais elle ne sont pas forcément très élaborées). Ainsi, l'utilisateur devra généralement prévoir des développements supplémentaires, et ce, même s'il n'envisage pas des études topologiques complexes. Ce sera le cas par exemple pour identifier les tronçons en impasse ou les boucles (fig. B.2.59).

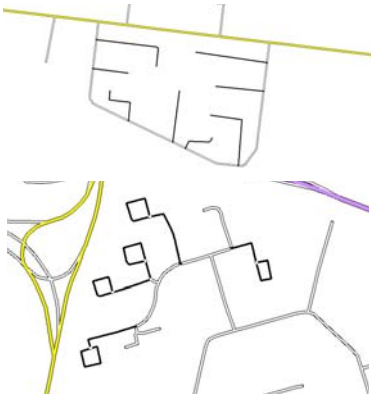


Figure B.2.59. *Caractérisation des 'îlots' en fonction des 'tronçons de route' internes*

Les caractéristiques topologiques traduisent des relations qui ne sont pas toujours quantifiables (ex. : attribut est/ n'est pas une impasse, est connecté à/ n'est pas connecté à). Au moment de l'agrégation elles donneront alors lieu à des indicateurs plus qualitatifs sous forme d'une caractérisation générale ou de classes avec leur effectif (par exemple : les tronçons sont principalement des impasses ou bien il y a 3 impasses et 1 boucle). Comme pour la composante géométrique, les analyses topologiques serviront à enrichir la description initiale des *objets géographiques* et à caractériser la répartition de l'agrégation dans son emprise (par exemple, position relative du 'bâti' par rapport aux espaces non construits fig. B.2.60, ou par rapport à d'autres objets cf. photographies p. 99). Les coûts et la difficulté liés à ces extensions seront comme pour les analyses morphologiques ceux inhérents aux développements informatiques.

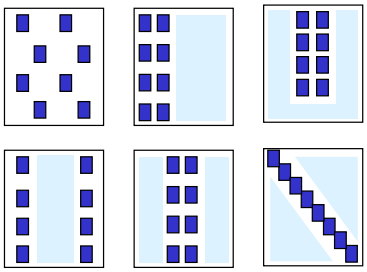


Figure B.2.60. *Caractérisation des 'îlots' en fonction de la dispersion de l''agrégation' dans son 'emprise'*

Les résultats des recherches effectuées en analyses à référence spatiale de type géométrique et/ou topologique, nous semblent particulièrement intéressants à exploiter pour développer et enrichir la méthode. Ils fournissent des indicateurs et des méthodes de complexités variables. Sous réserve de compétences informatiques, l'utilisateur intéressé pourra intégrer certains de ces résultats pour ajouter à son SIG quelques fonctionnalités (fig. B.2.61). L'investissement initial doit être envisagé en tenant compte du caractère générique de ces indicateurs ; ils peuvent s'appliquer à des objets thématiques différents.

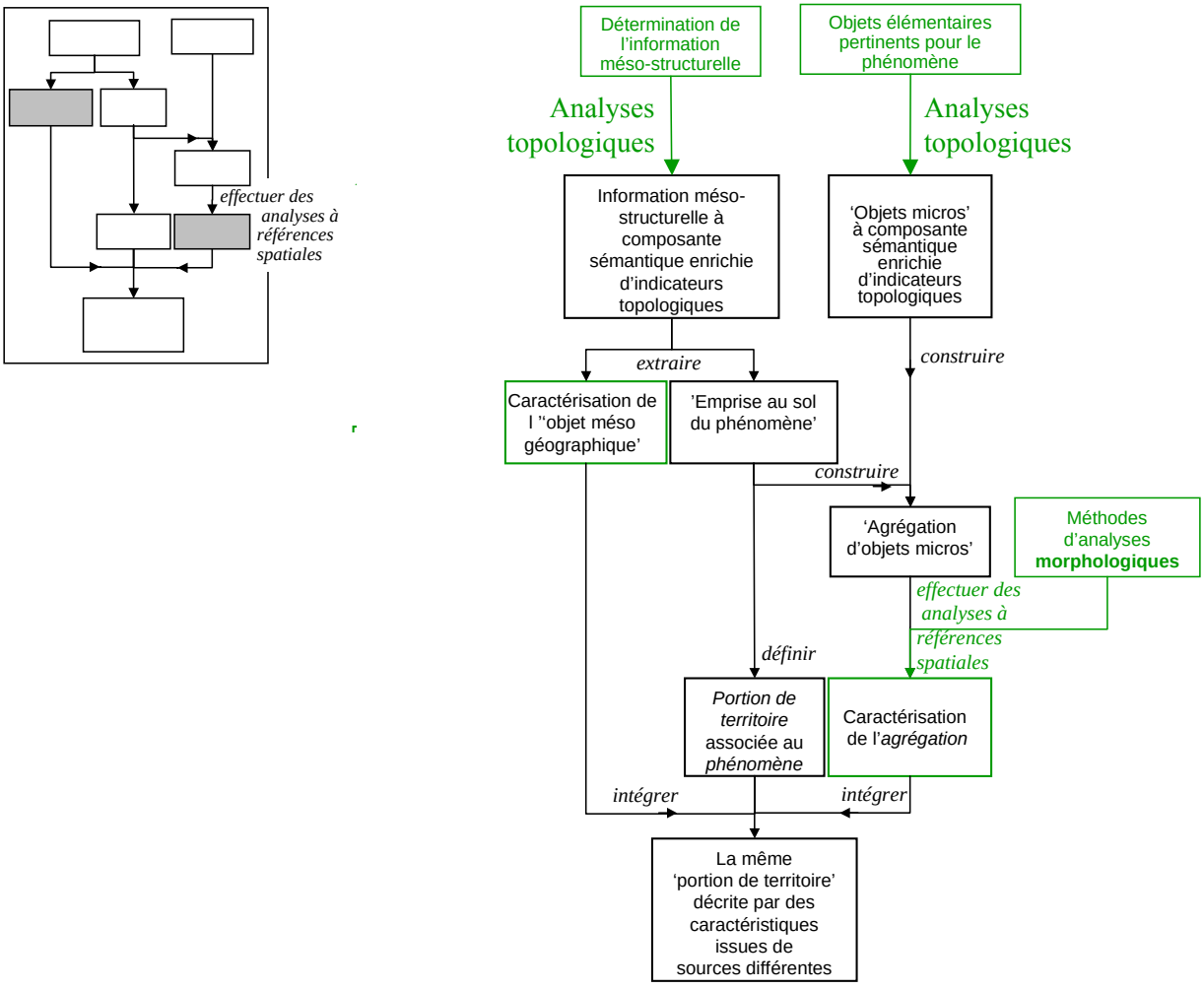

effectuer des analyses à références spatiales
Détermination de l'information méso-structurelle
Objets élémentaires pertinents pour le phénomène
Analyses topologiques
Analyses topologiques
Information méso-structurelle à composante sémantique enrichie d'indicateurs topologiques
'Objets micros' à composante sémantique enrichie d'indicateurs topologiques
extraire
construire
Caractérisation de l ''objet méso géographique'
'Emprise au sol du phénomène'
construire
'Agrégation d'objets micros'
Méthodes d'analyses morphologiques
effectuer des analyses à références spatiales
définir
Portion de territoire associée au phénomène
Caractérisation de l'agrégation
intégrer
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes

Figure B.2.61. *Enrichissement de la méthode par intégration d'analyses topologiques pour caractériser les 'objets' (méso et/ou micro) et les 'agrégations'*

B.2.6.4. Exploitation d'analyses statistiques

La composante sémantique peut également donner lieux à des études sophistiquées grâce à la mise en œuvre d'analyses statistiques élaborées.

L'utilisateur peut ainsi mettre à profit conjointement la méthode et une expertise en statistiques pour enrichir la description du phénomène étudié (fig. B.2.62). En particulier l'introduction de méthodes de classification pourra servir à identifier au sein de l''agrégation', des classes ou groupes d''objets' homogènes. Au sein des 'agrégations' même, on peut chercher à « regrouper » en classes des 'objets' identifiés comme «proches», pour mieux observer leur répartition dans l'espace et surtout par la suite pour tenter de saisir des évolutions différenciées.

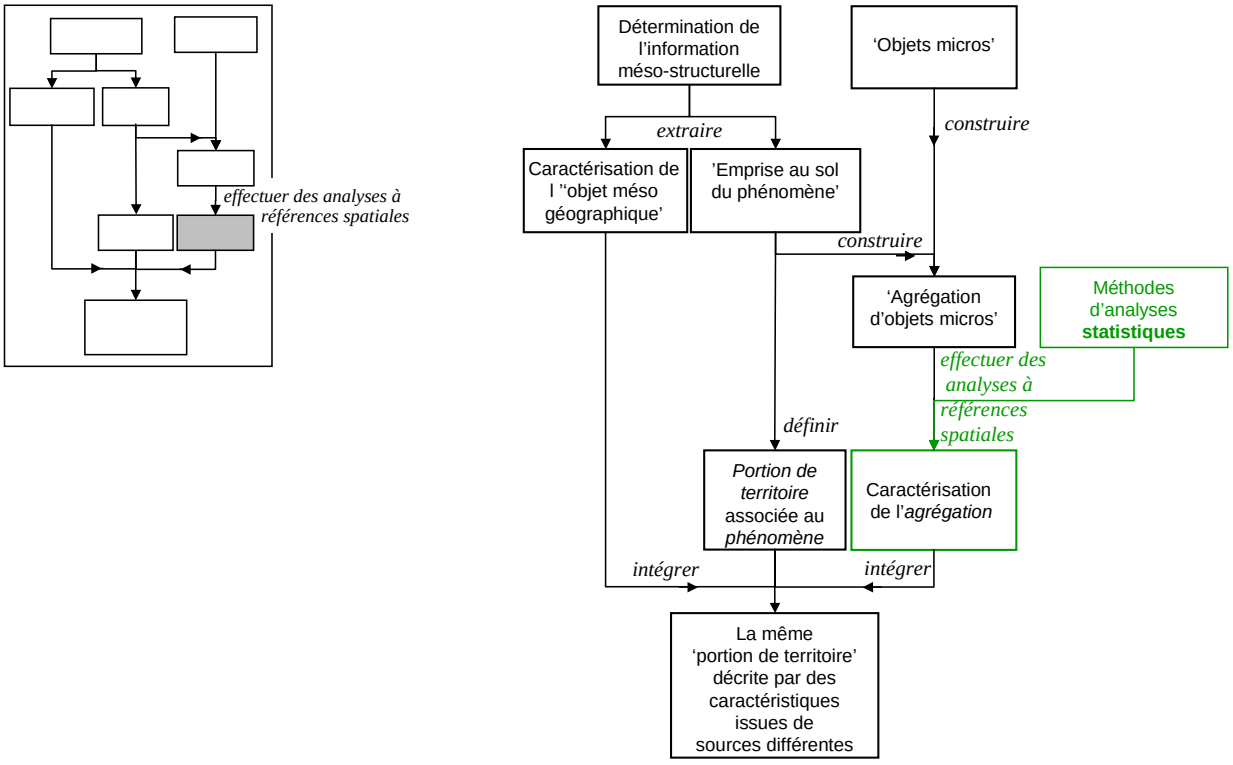


Figure B.2.62. *Enrichissement de la méthode par intégration d'analyses statistiques dans la caractérisation des 'agrégations'*

Par ailleurs, pour introduire un niveau d'observation supplémentaire, l'utilisateur pourra tenter de se placer délibérément dans le cas particulier (b2), pour pouvoir mettre en œuvre le méthode à plusieurs niveaux (cf. B.2.5.3.2). (Par exemple : il pourra introduire les 'îlots' comme objets d'études intermédiaires entre la 'zone pavillonnaire' observée et les 'bâtiments').

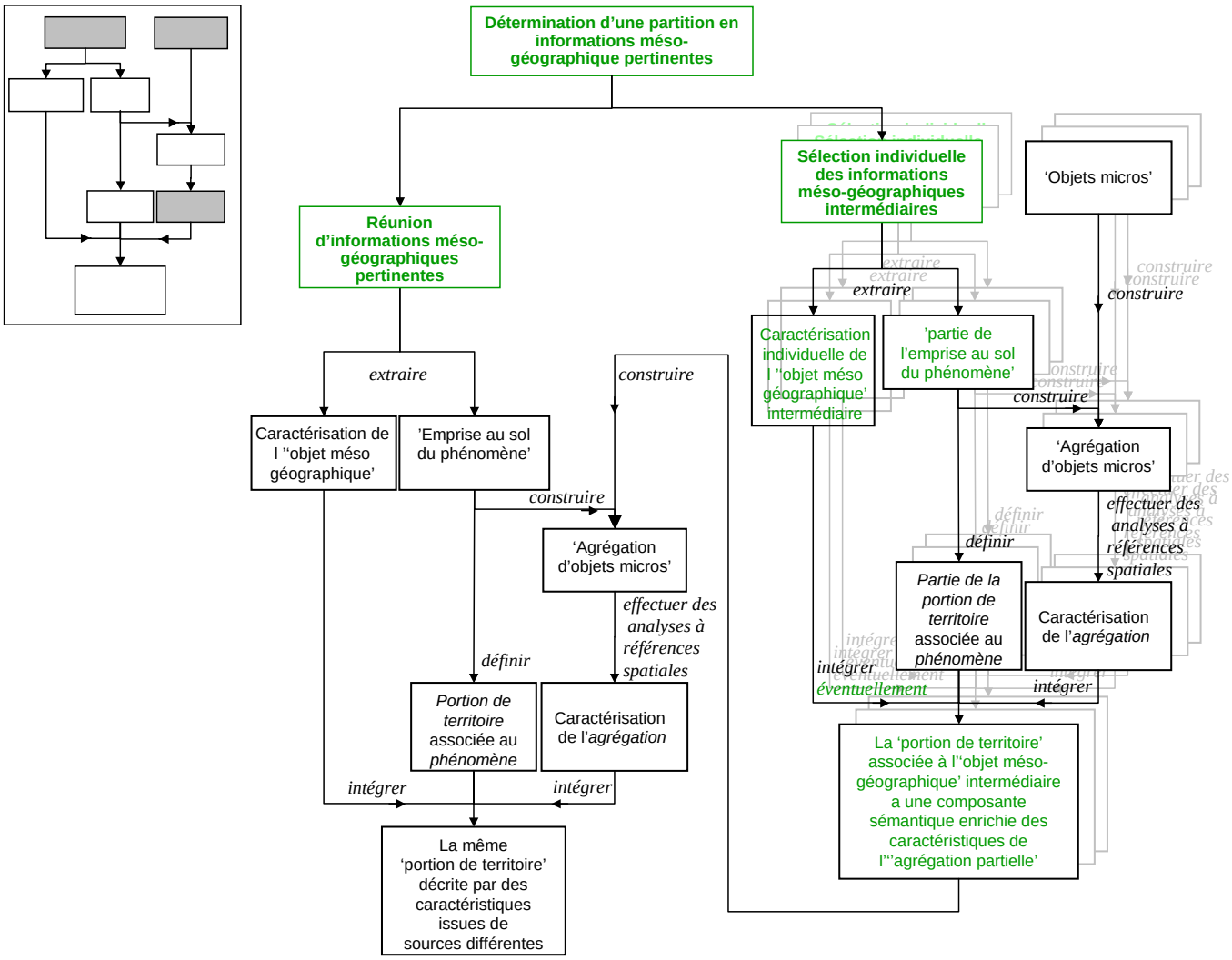

Détermination d'une partition en informations méso-géographique pertinentes
Réunion d'informations méso-géographiques pertinentes
Sélection individuelle des informations méso-géographiques intermédiaires
'Objets micros'
extraire
construire
Caractérisation de l ''objet méso géographique'
'Emprise au sol du phénomène'
'Agrégation d'objets micros'
définir
effectuer des analyses à références spatiales
Portion de territoire associée au phénomène
Caractérisation de l'agrégation
intégrer
La même 'portion de territoire' décrite par des caractéristiques issues de sources différentes
Caractérisation individuelle de l ''objet méso géographique' intermédiaire
'partie de l'emprise au sol du phénomène'
Partie de la portion de territoire associée au phénomène
éventuellement
La 'portion de territoire' associée à l''objet méso-géographique' intermédiaire a une composante sémantique enrichie des caractéristiques de l''agrégation partielle'

Figure B.2.63. *Enrichissement de la méthode par intégration d'un niveau d'observation supplémentaire*

L'introduction d''objets méso-structurels' intermédiaires permet de disposer à la fois d'informations synthétiques utiles au suivi général du phénomène et d'informations de détails sur le phénomène observé (fig. B.2.63). Ces informations seront particulièrement utiles dans la seconde phase de l'observation, lors du suivi, car elles permettront de saisir des évolutions du phénomène plus précises.

Outre les compétences en analyse spatiale nécessaires [Sanders, 2001], cet enrichissement soulève une difficulté formelle : il faut gérer plusieurs niveaux d'observations simultanément. Dans le SIG, elles résident surtout dans l'identification d''objets méso-structurels' pertinents. Les *îlots*, les *parcelles* (à plus grande échelle), les *communes* (à plus petite échelle) peuvent être de bons candidats. Ils sont :

- disponibles sous forme de données (condition nécessaire),
- surfaciques et forment une partition (pour permettre l'agrégation),
- généralistes (pour convenir à un grand nombre de phénomènes),
- enfin ils « structurent » l'espace géographique, car ils servent de référent pour la localisation de nombreuses informations; ils interviennent comme unité de collecte et comme unité d'analyse dans beaucoup d'études d'«acteurs du territoire ».

B.2.6.5. Utilisation de plusieurs agrégations

En introduisant la méthode d'agrégation par emprise et le concept d'objet méso-structurel nous nous somme assurés de la compatibilité des représentations entre 'agrégations' d'objets micro différents. Cette compatibilité qui permet de considérer simultanément plusieurs 'agrégations' dans une même étude, peut également servir à enrichir la description du phénomène (fig. B.2.64). Nous l'avons d'ailleurs testée dans le cas des *zones pavillonnaires* : nous avons construit des 'agrégations de bâtiments' *et* des 'agrégations de tronçons de routes'.

Le principe ici consiste à reproduire la méthode sur d'autres *objets élémentaires*. La difficulté principale réside dans l'identification d'objets micro' pertinents nouveaux. La caractérisation des nouvelles 'agrégations' pourra alors à son tour profiter des différentes techniques d'enrichissements précitées.

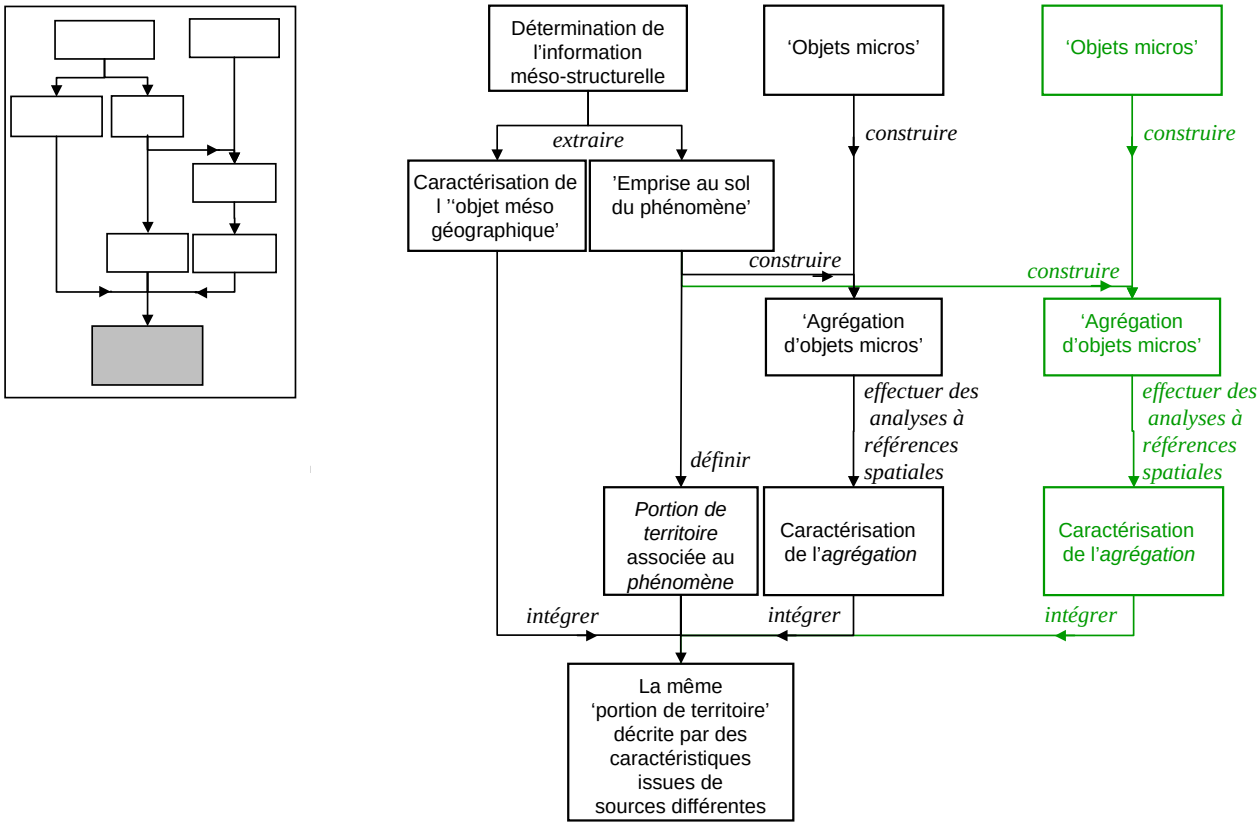


Figure B.2.64. *Enrichissement de la méthode par intégration d''agrégations d'objets micro' variées*

B.2.7. De l'"objet méso-structurel' à la 'portion de territoire'

La méthode d'agrégation par emprise proposée dans ce chapitre nous a servi à construire une méthode d'observation multi-niveaux. La définition d'un 'objet méso-structurel' nous a permis de l'adapter aux cas où l''emprise' servant d' «emporte-pièce» n'est pas donnée a priori ; c'est à dire quand il n'existe pas de données représentant l'*objet d'observation principal* pour fournir directement un «emporte-pièce» adéquat, grâce à sa composante géométrique (fig. B.2.65a).

Si l'on reprend la définition de l''objet méso-structurel', c'est un objet technique possédant deux propriétés :

- il existe des données qui disposent d'une composante géométrique surfacique propre pour le représenter (fig. B.2.65b)
- et la 'portion de territoire' caractérisée par cette composante géométrique peut être considérée comme une représentation de l'*emprise au sol du phénomène*.

Au travers de la notion de «portion de territoire», l''objet méso-structurel' établit un lien entre l'*objet d'observation principal* qui n'a pas de représentation en données et des données ayant une composante géométrique. L'*objet* méso-structurel qui a été introduit comme un 'objet' méthodologique, nous semble en fait être plus que cela ; il peut être identifié à un 'objet géographique' athématique : 'portion de territoire'. Nous avons déjà constaté qu'il pouvait correspondre à un 'objet géographique' particulier (cf. p. 140) ; cela permet alors d'enrichir les caractéristiques observées avec des caractéristiques qui lui sont propres (fig. B.2.65c). Nous avons vu également que, lorsque l'information méso-structurelle se présente sous forme d'une partition, les 'objets méso-structurels' qui contribuent à cette partition peuvent jouer le rôle d'objets d'observation secondaire dans une observation conçue à plusieurs niveaux.

L'**introduction des objets géographiques *portions de territoire*** nous paraît **introduire une articulation conceptuelle** supplémentaire dans la modélisation des entités géographiques en données géographiques (fig. B.2.65d), qui présente une analogie avec celle introduite par les SIG par rapport aux outils de cartographie numérique :

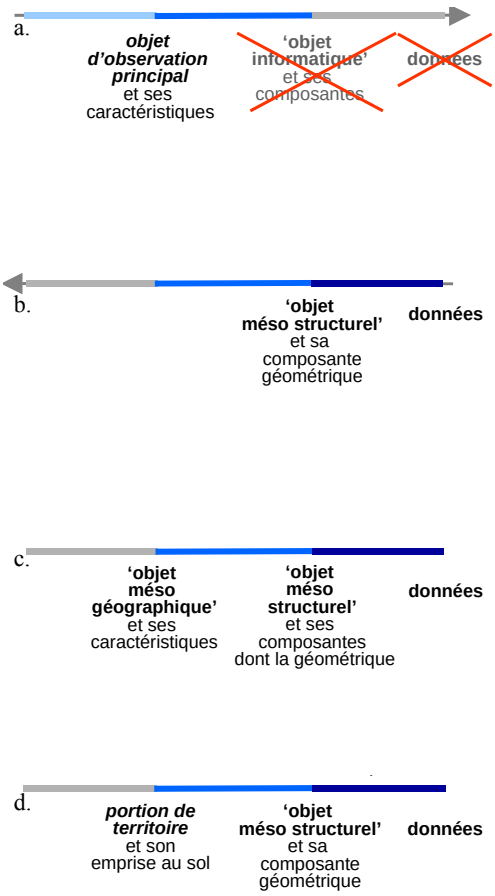


Figure B.2. 65. *L' 'objet méso-structurel' et les étapes de formalisation*

L'apport conceptuel des SIG par rapport aux outils de cartographie numérique.

Les outils de cartographie numérique associent les caractéristiques graphiques directement aux primitives géométriques. Ainsi, si l'on veut établir une carte du réseau routier en fonction du classement fonctionnel, il faut sélectionner manuellement toutes les polylignes représentant une autoroute et leur associer les différentes caractéristiques graphiques correspondant à la représentation désirée (couleur, symbole, taille, etc.), puis faire de même pour les nationales, les départementales, etc. Si l'on veut ensuite établir une carte du réseau routier en fonction du nombre de voies, il faut alors tout recommencer : sélectionner manuellement toutes les polylignes correspondant aux routes à une voie et leur associer les caractéristiques graphiques souhaitées et recommencer pour celles à deux voies, et ainsi de suite. L'informatisation permet d'automatiser seulement l'édition répétée d'une même carte, mais le travail manuel (de sélection) reste important.

En introduisant la composante sémantique comme intermédiaire entre la composante géométrique et la « composante graphique », les SIG a considérablement assoupli le processus de rédaction cartographique. En effet, il suffit d'associer une fois les caractéristiques graphiques aux valeurs d'attributs correspondants. Le SIG utilise alors cette légende pour représenter automatiquement de la bonne façon les primitives des différents 'objets' en fonction de leur composante sémantique. Si L'utilisateur veut ensuite changer de carte, il lui suffit de définir une autre légende.

L'introduction de l''objet portion de territoire' entre l''objet géographique' et sa représentation sous forme de données nous paraît apporter une souplesse du même genre.

Formellement, au lieu de décrire directement chaque 'objet géographique (par nature)' par une composante géométrique et une composante sémantique propre (fig. B.2.66), on pourrait introduire un objet 'portion de territoire' : à chaque 'objet géographique' seraient alors associés un objet 'portion de territoire' et une composante sémantique, la composante géométrique étant associée uniquement à la 'portion de territoire' (fig. B.2.67). Les 'objets géographiques' deviendraient alors des caractéristiques de la 'portion de territoire', des informations localisées (cf. p. 106) ayant pour support la 'portion de territoire'.

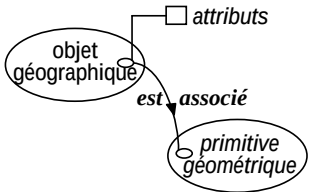


Figure B.2.66. *Modélisation classique dans les SIG*

Ainsi, des 'objets' qui partageraient la même 'portion de territoire' seraient assurés de partager la même composante géométrique. Leur représentation serait ainsi compatible, ce qui réduirait certains problèmes de représentations multiples (fig. B.2.68). Sur la base d'une 'portion de territoire' commune, on pourraient transférer des caractéristiques d'un de ces 'objets' vers un autre, chaque 'objet' étant alors considéré comme la même 'portion de territoire' vue d'un point de vue thématique particulier (cf. p. 94) (par exemple si la même 'portion de territoire' est support à la fois d'une unité de recensement, d'une zone de PLU, d'une zone topographique, l'urbaniste pourra intégrer l'ensemble des caractéristiques pour disposer d'une vision plus riche de cette *unité urbaine*).

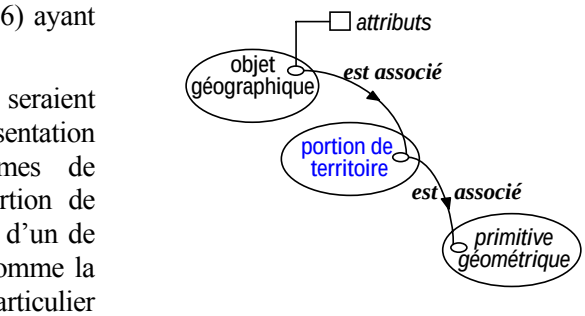


Figure B.2.67. *La portion de territoire entre 'objet' et primitive géométrique*

Tel est conceptuellement le principe mis en œuvre dans la méthode d'observation multi-niveaux.

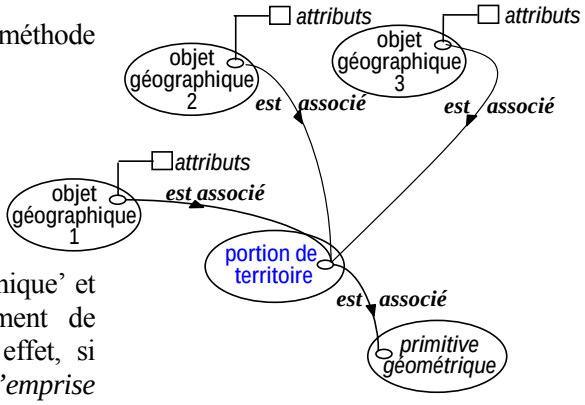


Figure B.2.68. *La portion de territoire support des objets géographiques*

Le fait d'introduire la 'portion de territoire' entre l''objet géographique' et sa composante géométrique apporterait par ailleurs un élément de souplesse supplémentaire dans la formalisation technique. En effet, si l'introduction de l'objet 'portion de territoire' pour modéliser *l'emprise* d'un *objet géographique* permet de distinguer la 'portion de territoire' de ses diverses représentations en information géographique, elle permet également par ailleurs de suivre une même *portion de territoire* au travers de représentation en données différentes (fig. B.2.69).

Nous avons vu qu'un même 'objet' lorsqu'il est numérisé par deux personnes différentes a toutes les chances d'avoir une composante géométrique différente. Ceci est encore plus vrai lorsque les représentations de cet 'objet' ne participent pas des mêmes bases de données (fig. B.2.2). Ainsi, l'*îlot* défini comme une maille du réseau routier aura t'il une représentation différente dans la BD TOPO® et dans BASE ILOT® de l'INSEE. Au moment d'intégrer les données issues des deux bases, il en résulte alors des incohérences (cf. figure B.2.70) (qui font d'ailleurs partie des problèmes de gestion de représentations multiples dans les SIG). Ainsi, si formellement il s'agit bien du même objet *îlot*, et même si les spécifications sont proches et les systèmes de références géodésiques égaux, il peut y avoir des *bâtiments* appartenant à l'*îlot* qui soient représentés comme « à cheval » où en dehors, car leur représentation ne fait pas référence aux mêmes données 'îlot' support (cf. figure B.2.71).

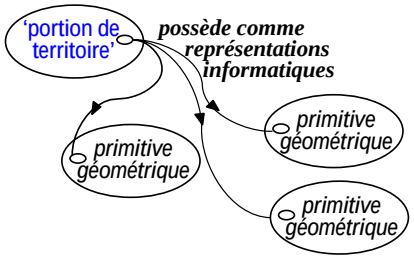


Figure B.2.69. *La portion de territoire est ses différentes représentations en données géographiques*

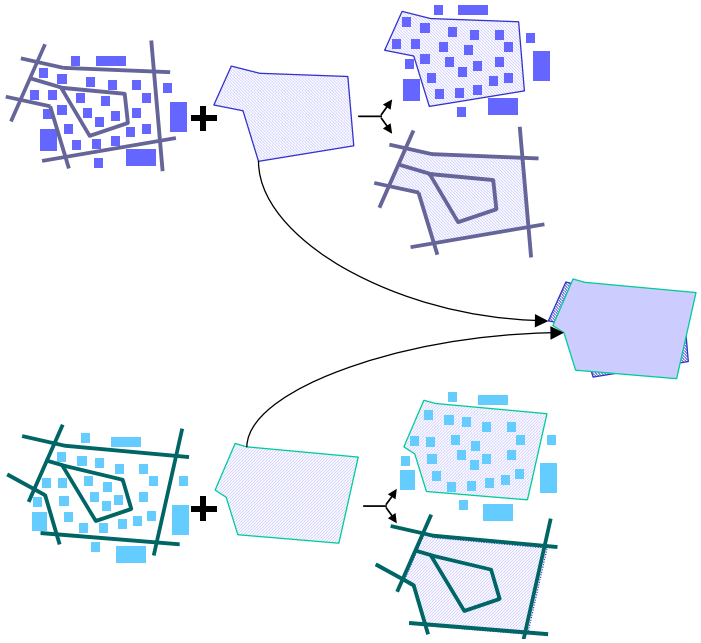


Figure B.2.70. *Des données représentant un même objet', mais issues de bases différentes ont une forte probabilité d'être géométriquement incohérentes.*

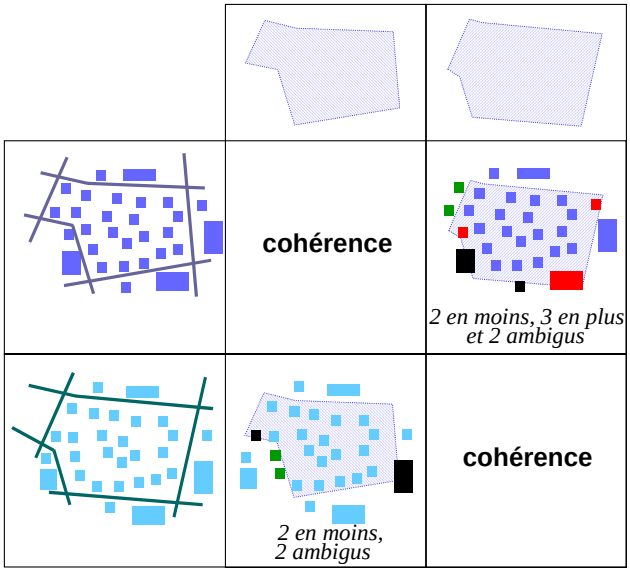


Figure B.2.71. *Le choix de la donnée utilisée comme emporte-pièce pour réaliser les 'agrégations d'objets micro' influe sur les résultats de ces agrégations*

Le fait d'introduire la *portion de territoire* permet :

- de justifier à un niveau formel, le transfert des caractéristiques des agrégations d''objets micro' vers un même 'objet' : il s'agit d'une mise en commun des caractéristiques des *objets*, sur la base d'une *portion de territoire commune*,

- de distinguer la '*portion de territoire'* de ses représentations numériques. Il permet ainsi de travailler avec plusieurs bases de données simultanément, en laissant la possibilité de choisir la composante géométrique la mieux adaptée pour chaque agrégation d''objets micro' (en particulier celle qui

conserve le maximum de cohérence topologique avec les 'objets micro' à agréger) tout en maintenant la compatibilité des emprises nécessaires au transfert de caractéristiques.

- de disposer au final, d'indicateurs plus cohérents et plus précis.

Or nous allons maintenant voir, que le fait de saisir l'apparition ou la disparition d'un 'objet micro' (ex. : création d'un 'bâtiment' dans un 'îlot') peut avoir son importance…

B.3. Suivi de phénomènes géographiques avec un SIG

B.3.1. Suivi d'évolutions de phénomènes géographiques

Le suivi d'un phénomène géographique avec un SIG nécessite d'introduire la dimension temporelle dans l'observation. Il existe des solutions techniques pour suivre des informations géographiques dans le temps ; nous les avons présentées dans le chapitre A.3. Il s'agit soit de modèles de gestion d'informations temporelles conçus préalablement à la constitution des données, soit d'outils d'appariement et de calcul de différentiels permettant d'extraire des mutations. Cependant ces solutions travaillent sur des données ; les informations sur les évolutions qu'elles produisent portent sur des 'objets informatiques', dont on a vu dans le chapitre B.1 qu'ils ne se confondent pas avec les *objets géographiques* étudiés par le thématicien. Les informations d'évolutions fournies aux utilisateurs sont de fait, des informations pré-interprétées (ex. : un 'bâtiment' peut être déclaré détruit alors que l'*objet géographique bâtiment* correspondant n'a subi que des modifications architecturales (réaménagement) sans aucun changement de fonction, de propriétaire, d'activité, etc.). Par ailleurs ces solutions requièrent une codification (dans la spécification des modèles temporels ou dans la définition des informations à extraire), ce qui a pour effet de rigidifier encore le cadre d'interprétation des évolutions pour le thématicien.

Dans ce chapitre, nous montrons que les informations élaborées avec la méthode proposée ne nécessitent pas de solution technique complexe pour saisir leurs évolutions : l'extraction automatique des mutations et leur interprétation peuvent être effectuées simplement, par le thématicien, au niveau des *objets géographiques*. L'utilisateur peut ainsi déterminer lui-même si les *objets* changent, se transforment et/ou disparaissent. Il pourra l'établir avec d'autant plus de précision que la méthode mettra à sa disposition une caractérisation plus riche.

B.3.1.1. Evolution et changement

B.3.1.1.1. « L'objet suffisamment lui-même »

D'après le dictionnaire [Petit Robert, 2001] une évolution est : « *une suite de transformations dans un même sens ; une transformation graduelle assez lente ou formée de changements successifs insensibles* », tandis qu'un changement est *« le fait de ne plus être le même »*. L'évolution comporte donc une notion de *continuité* de l'objet qui reste alors « encore lui-même » malgré les transformations ou changements insensibles, tandis que le changement établit une distinction avec un *état d'avant* et un *état d'après*. Or il est parfois délicat de déterminer le moment où un objet est « encore lui-même » et quand il cesse de l'être. Par exemple, dans le phénomène de mitage quand doit-on considérer qu'une parcelle d'espace vert a été « colonisée » par de l'habitat : Quand 5% de sa surface est construit, quand la densité de population est supérieure à 0,05/Km^2, quand il y a plus de 3 familles qui y résident en permanence, quand il y a plus de 5 constructions temporaires ?… La figure B.3.1 illustre cet embarras à définir où se situe la limite exacte entre *évolution* et *changement.*

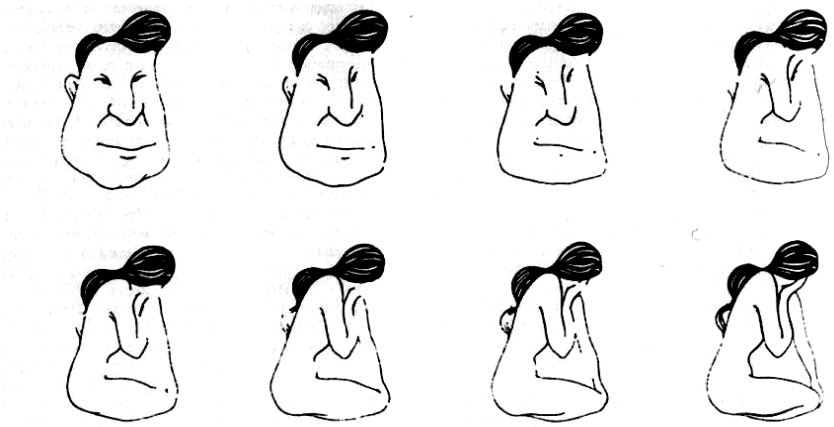


Figure B.3.1. *Quand le visage devient-il silhouette ?*
(Dessin de F. Attneave, extrait de [Reuchlin, 1990, page 560])

En nous inspirant de l'utilisation de l'adverbe « suffisamment » par Winnicott [David et Wallbridge, 1992, p181], nous proposons de définir la notion *« d'objet suffisamment lui-même »*. Cette expression nous permet d'évoquer *l'objet* que l'on peut considérer dans ses caractéristiques principales comme assez proche de ce qu'il était avant le(s) changement(s) pour être tenu comme semblable. Ainsi, nous posons :

Objet suffisamment lui-même : *Un objet est suffisamment lui-même si ses caractéristiques évoluent dans des proportions telles que, pour une observation donnée, l'objet conserve son « identité ». Toutefois, pour un autre observateur, cet objet peut avoir suffisamment changé pour être devenu « autre ». La problématique de l'objet suffisamment lui-même comporte une dimension temporelle. En effet, un objet peut être suffisamment lui-même entre plusieurs dates d'observation, mais ne pas l'être entre l'état initial et l'état final de l'observation.*

Cette expression nous sert à pointer la complexité d'un suivi d'évolution. Elle permet de rappeler la limite parfois ténue entre la « permanence » de l'*objet* (le temps où il continue à être ce qu'il est) et le moment de la « rupture » (où il disparaît ou devient *autre*) dans ce que l'on pourrait appeler la « continuité »[1]. Par exemple la figure B.3.1 qui donne à voir l'évolution de l'objet-visage (sous forme d'un « film » constitué par les différentes observations) illustre combien il est délicat de répondre à la question du *changement* (en l'occurrence le changement du visage en

1. {Esposito, 2006} utilise ces trois termes en architecture :
« La *rupture* est définie comme une déstructuration d'une forme existante, qui peut aboutir à de nouvelles recompositions
La *permanence* correspond à la résistance d'un élément aux mutations, qui témoigne de son enracinement dans le territoire sur la longue durée,
La *continuité* est une transformation qui se met en relation avec le patrimoine à travers sa récupération et/ou sa ré interprétation créative ».

silhouette). En effet, à quel moment l'objet-visage n'est-il plus un visage ? Est-il alors déjà devenu l'objet-silhouette dans un changement ou bien y a t'il une ambiguïté (avec l'existence d'un objet à la fois visage *et* silhouette) ou encore y a t'il une forme intermédiaire qui ne soit ni visage ni silhouette ? Ces questions portent ainsi sur la forme du changement lui-même : s'est-il opéré avec une rupture dans le cheminement ? ou bien est-il une évolution qui de façon graduelle a fait émerger un nouvel objet , en passant ou non par une forme indéterminée? Enfin, l'objet final apparemment très différent de l'objet initial peut-il être considéré malgré tout comme une évolution de ce dernier ? (fig. B.3.2)... Nous ne souhaitons pas entrer dans des questions philosophiques, ni dans des modèles complexes (ou un même objet pourrait être deux objets différents ou bien un *non-objet*). En revanche, nous voulons dégager ici l'importance d'une réflexion préalable sur les propriétés essentielles des *objets* observés, sur la nécessité d'identifier le plus précisément possible ce qui les caractérise, ainsi que les conditions et limites en dehors desquelles l'*objet* cesse d'être ce qu'il est. Ainsi, relativement au phénomène de mitage, on pourrait caractériser les 'parcelles' d'espace vert sur la base du 'bâti' en distinguant par exemple trois types de 'parcelles' :

- les parcelles « vulnérables » : celles au sein desquelles il existe au moins k constructions temporaires ou une construction en dur,
- les parcelles « sensibles » : celles qui accueillent au moins k' constructions en dur,
- les parcelles « en devenir urbain » : celles dont la densité de construction est supérieure à 0,02[2], qui ne sont déjà « plus vraiment » des espaces verts et que l'aménageur cherchera à organiser au mieux en espaces urbanisés.

2. La valeur 0,02 est une valeur que l'on peut trouver dans certains PLU ([www.lacub.com]).
Quant aux valeurs k et k', elles seront définies par le thématicien en fonction par exemple de la taille de la parcelle.

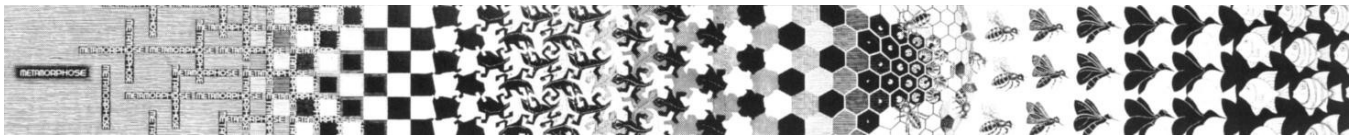


Figure B.3.2. *Une évolution est une suite de transformations graduelles qui peut aboutir à une métamorphose, à un changement radical (tableau de M.C. Escher, La métamorphose II, 1940, extrait de [Ernst, 1994])*

B.3.1.1.2. Caractéristiques d'une observation temporelle

La caractérisation des *objets géographiques* prépare la définition des composantes des 'objets informatiques'. Dans la perspective d'un suivi avec un SIG, il est important que la formalisation technique puisse intégrer au mieux cette caractérisation. Cette étape constitue en effet l'étude préalable à la modélisation en données géographiques. Elle assure à une information de pouvoir être observée directement dans le SIG (cf. chapitre B.1). Il s'agit ici de saisir les informations essentielles, mais aussi les détails utiles au suivi des *objets* étudiés. La caractérisation doit donc être faite avec des *niveaux de détails* suffisants pour saisir des *changements insensibles* et pas seulement des transformations ou des changements. Toutefois, pour observer des évolutions, il ne suffit pas de disposer de niveaux de détails importants. Il faut également que les paramètres temporels de l'observation (pas et date à l'origine) soient adaptés. En effet, par définition, l'évolution est *graduelle*. Cependant, entre l'état initialement considéré et l'état final, l'objet peut être devenu totalement autre. La figure B.3.3 illustre ainsi deux états très différents pouvant correspondre aux états avant/après un changement. Rien ne permet de savoir s'il s'agit d'une évolution lente ou si la situation est apparue brusquement. Pour cela, il aurait fallu disposer d'observations intermédiaires, d'un *pas d'observation* plus fin.

Figure B.3.3. *Entre l'état initial et l'état final, il peut s'être opéré un véritable changement*

Le *pas d'observation* est l'intervalle de temps séparant deux observations. Il peut être choisi régulier (par exemple tous les quatre ans, tous les ans) ou irrégulier. Le choix du pas dépend du phénomène étudié et de son rythme d'évolution[3]. Il s'agit de disposer d'une « sensibilité » aux changements suffisante pour les saisir ; l'utilisateur doit pouvoir discriminer les changements en évitant néanmoins des observations inutiles et redondantes (généralement coûteuses).

3. Il dépend également du type d'observation. (Par exemple, pour des observations rétrospectives, on pourra choisir un pas irrégulier correspondant à la survenue d'événements marquants).

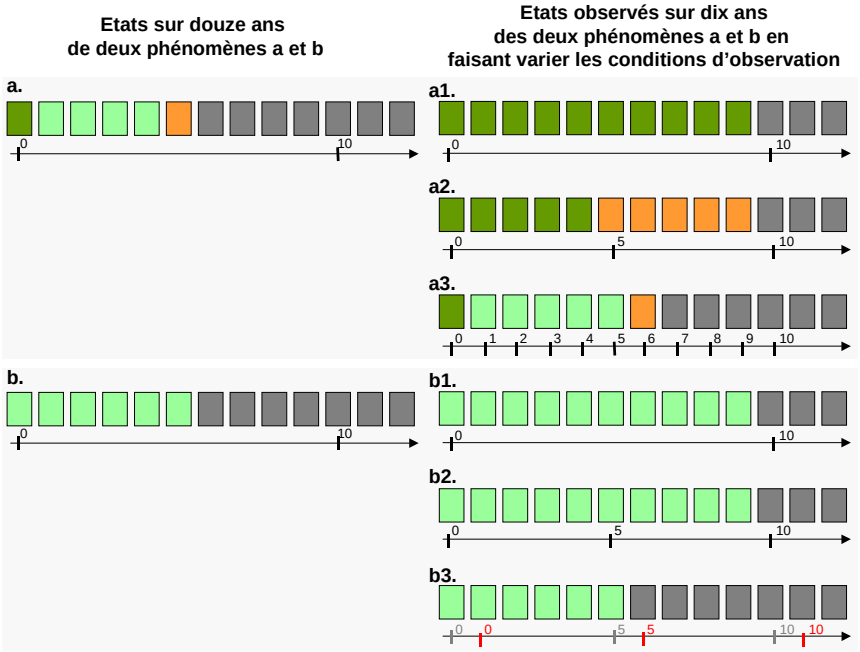


Figure B.3.4. *Phénomènes et conditions d'observations*

La première colonne montre deux phénomènes a et b, dont elle présente douze états consécutifs. La seconde colonne montre ce que l'on observe en faisant varier les conditions d'observation. Ainsi, sur le phénomène a, nous constatons les conséquences d'un pas mal adapté (de 10 et de 5 ans respectivement pour a1 et a2). Si l'on choisit d'effectuer une observation avec un pas régulier, alors pour saisir les changements entre la cinquième et la septième année un pas annuel (a3) devient nécessaire. Cette régularité coûteuse qui est nécessaire en cours d'évolution n'est cependant plus utile dans la phase de stabilité. Sur le phénomène b, nous constatons également l'influence de la date d'origine de l'observation. En débutant l'observation à la date 0, on note que ni le pas de 10 ans (b1), ni celui de 5 ans (b2) ne permettent de déceler le changement intermédiaire. En revanche avec une autre date d'origine, le pas de 5 ans peut alors convenir (b3).

Or il est difficile de connaître le rythme des changements des objets observés et les dates de leur survenue[4]. Ainsi, même en choisissant un pas variable pour s'adapter au mieux au phénomène, les observations risquent d'être insatisfaisantes, car « en retard » sur ce qui se produit sur le terrain (cf. figure B.3.5).

4. Il faudrait pouvoir prévoir pour mieux observer...

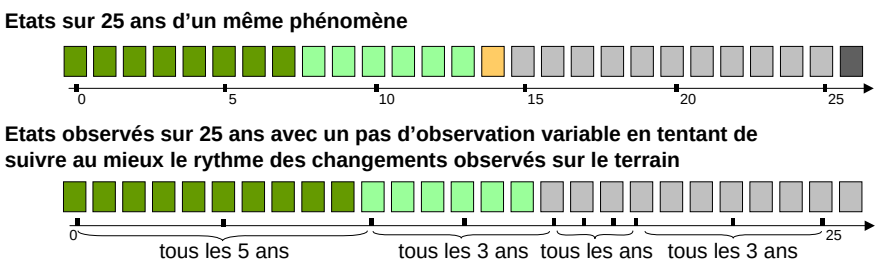

Etats sur 25 ans d'un même phénomène
Etats observés sur 25 ans avec un pas d'observation variable en tentant de
suivre au mieux le rythme des changements observés sur le terrain
tous les 5 ans
tous les 3 ans
tous les ans
tous les 3 ans

Figure B.3.5. *De la difficulté d'adapter en temps réel le pas d'observation aux évolutions d'un phénomène*

Le choix du *pas d'observation* dépend également du caractère fini du phénomène, de la durée totale de l'observation envisagée. De fait, il finit souvent par être un compromis entre des arguments thématiques en faveur d'un pas fin pour ne « rien rater » et des arguments techniques et économiques (relatifs au temps irréductible nécessaire à la collecte et aux traitements des informations et aux coûts de ceux-ci) qui poussent à choisir des intervalles de temps plus « épais ».

Une autre caractéristique temporelle de l'observation doit également être prise en compte : *la date à l'origine* de l'observation, c'est à dire la date à laquelle débute le suivi (cf. figure B.3.4.b3). Pour observer un phénomène ou ses effets, (par exemple, pour observer la mise en place d'une nouvelle urbanisation ou ses effets, ou encore les effets d'une nouvelle réglementation), il est préférable de débuter les observations avant que les premiers changements n'aient commencé. On notera ici que le caractère encore récent des bases de données numériques peut à ce titre limiter l'observation d'un phénomène avec un SIG.

B.3.1.2. Suivi de phénomènes géographiques dans un SIG

B.3.1.2.1. Un « décalage » entre évolutions d'objet et changements de donnée

L'observation cinématique d'un phénomène géographique implique de pouvoir suivre dans le temps les évolutions des objets sur lesquels le phénomène se produit. Nous avons vu dans le chapitre A.3, que les applications SIG comportent généralement peu d'informations temporelles. Lorsqu'il en existe, elles concernent alors des changements sur les données et non des évolutions d'objets. Or les changements ne sont pas des évolutions. Et les données ne sont pas les objets ; elles n'en sont que des représentations déformées par le filtre des spécifications, qui ne donnent à voir que les informations qu'elles contiennent (cf. chapitre B.1).

L'observation des changements sur les données n'équivaut donc pas à celle des évolutions d'objets géographiques. Proposer un suivi d'évolution de phénomènes géographiques à l'aide d'un SIG implique donc d'aider l'observateur à réduire ce « décalage », en lui permettant de tirer de l'observation dans un SIG des informations utiles à l'analyse des objets qu'il étudie. On peut pour cela proposer une solution technique consistant :

– d'une part à enrichir les données en en acquérant de nouvelles, pour obtenir une représentation au « plus près de la réalité » ;

– et d'autre part à proposer des modèles et moyens techniques sophistiqués d'analyses temporelles .

B.3.1.2.2. Une approche technique

Une telle solution risque toutefois d'être coûteuse à mettre en œuvre pour l'utilisateur, tant économiquement (lorsqu'il s'agit d'acquérir de nouvelles données), que pratiquement (puisque cela lui demande de se familiariser suffisamment avec la gestion de la dimension temporelle pour pouvoir implémenter un modèle qui lui convienne). Pour alléger ce travail, les producteurs de données génériques réfléchissent aux outils et méthodes pour faciliter le passage de relais entre la production des données et leur exploitation chez les utilisateurs[5]. Ainsi, [Badard, 2000, p.49] intègre dans ses travaux une typologie des changements détaillée qui s'inspire de typologies d'évolutions plus géographiques [Cheylan, 1994], [Claramunt, 1994] qui *« tend à traduire des évolutions plus proches de la réalité terrain, c'est à dire plus proches de phénomènes observables »*. Il introduit en particulier la *scission*, l'*agrégation* et la *fusion* (cf. A.3

5. Les producteurs de données génériques produisent des données dites *support* ou *de référence*. Ces données servent aux utilisateurs à localiser leurs propres données thématiques, soit par géocodage (ex. : géocodage des entreprises à l'adresse), soit par numérisation (ex. : numérisation d'une nouvelle zone manuellement ou à l'aide d'outils tels que les buffers), soit en l'intégrant à une composante sémantique existante comme attribut (ex. : enrichissement des informations descriptives sur une commune).
Aujourd'hui, les producteurs de données support cherchent en premier lieu à faciliter la mise à jour des données des utilisateurs. En effet, lorsqu'un utilisateur utilise les données d'un producteur pour localiser ses propres données, au moment des mises à jour des données support, il rencontre souvent des difficultés à reporter automatiquement ses données dans la nouvelle version de la base. Parfois, il lui faut même refaire entièrement le travail d'intégration. Ce travail étant fastidieux, coûteux et sans valeur ajoutée, la demande première des utilisateurs portent aujourd'hui essentiellement sur la fourniture de moyens de reports automatiques de leurs données. Cette demande se formule en particulier par l'attente d'un identifiant pour enrichir le modèle par *snapshots* qu'ils exploitent couramment.
Dans les faits, pour les utilisateurs, il s'agit aujourd'hui moins de faire du suivi temporel, que de pouvoir déjà disposer à moindre coût d'un état des lieux à jour de leurs propres informations.

note 10). Toutefois cette classification qui convient bien aux évolutions de parcelles cadastrales peut être moins bien adaptée aux évolutions d'autres objets géographiques. Plus généralement, elle soulève la question de la ***pré-interprétation des changements de données.***

D'un point de vue technique, en toute rigueur, on devrait considérer avec [Bédard, 1997] qu'il n'y a que deux types de changements effectifs sur les données : la *création* et la *destruction*, auxquels on peut adjoindre le « non-changement » que l'on appelle généralement l'*identité* (au sens de à l'identique). L'exploitation d'une clé MD5 (qui associe automatiquement un identifiant unique à toute donnée à partir de ses éléments caractéristiques, cf. p. 83) va d'ailleurs en ce sens. En effet, toute modification de la donnée, quelque soit sa forme et son ampleur, entraîne un changement systématique d'identifiant (de *signature*) : la donnée numérique en tant que configuration de bits spécifique n'est alors « plus la même ». Cependant, une telle typologie ne permet pas de distinguer des situations géographiques très variées. Par exemple, une extension d'un bâtiment, une modification d'affectation, une modification du propriétaire, une correction d'erreur quant à la date de construction, etc. correspondront toutes à la *destruction* du 'bâtiment' suivie d'une *création* d'un nouveau 'bâtiment'. Au niveau des *objets géographiques*, il sera impossible d'interpréter automatiquement l'information *destruction* pour discerner les différents cas. De plus, si aucun lien n'existe entre le 'bâtiment détruit' et le 'bâtiment créé', alors au niveau du *bâtiment (géographique)*, cette destruction pourra être interprétée comme une destruction pure et simple de celui-ci.

Cependant techniquement, il est possible d'introduire des types de changement plus subtils (ex. : modification géométrique ou sémantique). Il suffit pour cela d'assimiler les données aux informations sur les composantes de l''objet informatique'. Par exemple, les changements opérés sur un polygone représentant la composante géométrique d'un 'objet informatique' sont interprétés comme une modification géométrique de cette composante et les changements de valeur d'attribut sont interprétés comme des modifications sémantiques. Ces interprétations nécessitent déjà:
- de pouvoir suivre les différentes représentations d'un 'objet informatique' en données ; il faut donc à la fois un lien pour signifier que des données représentent la même information (par exemple un identifiant) et que ce lien soit temporel pour un suivi des représentations dans le temps,
- de pouvoir déterminer les mutations entre deux données.

Autrement dit pour proposer cette typologie, le producteur de données doit mettre en place sur la base de données qu'il produit des solutions élaborées de suivi des évolutions.

Par ailleurs cette interprétation implique des spécifications. Par exemple, *« si le changement opère sur des attributs « secondaires » (ex. : changement du nombre d'habitants), le changement sera interprété comme une modification sémantique », « si le changement se produit sur des attributs principaux alors l''objet informatique' sera considéré comme détruit (ex. : si la fonction du bâtiment change) »*. De même, *« si la primitive géométrique est modifiée « très partiellement », on considérera que l''objet informatique' a subi une modification géométrique ». « Si la modification est plus importante alors il sera considéré comme détruit »*. Ces spécifications introduisent au niveau des 'objets informatiques' une pré-interprétation des évolutions. Or, la problématique de l'objet suffisamment lui-même a montré la difficulté de définir quand un *objet géographique* est modifié, quand il n'est plus le même. Par exemple que décider si l'objet subi un changement de la totalité de ses attributs secondaires ? ou encore si la totalité du périmètre est modifiée, mais avec un écart de moins de 10 cm ? L''objet informatique' doit-il être considéré

comme détruit ou comme modifié ? Un autre exemple est donné par la figure B.3.6 où l'on voit un bâtiment en grande partie détruit (seule sa façade sur rue a été conservée). Les changements opérés sur la primitive géométrique ont de grandes chances d'être interprétés comme une destruction du 'bâtiment' correspondant. Or, en gardant cette façade ce bâtiment reste suffisamment lui-même, pour que d'un point de vue légal, il puisse conserver le coefficient d'occupation des sols associé au bâtiment initial, alors qu'une construction nouvelle aurait dû respecter un coefficient moindre. Ainsi, d'un certain point de vue, ce bâtiment est resté le « même », il n'a pas été détruit.

Figure B.3.6. *Modification géométrique ou destruction / création ?*

Sauf spécifications particulières, le fait de ne conserver que la façade sur rue d'un' bâtiment' sera interprété comme une destruction du 'bâtiment' pour la création d'un nouveau. Cependant suivant le point de vue thématique, l'interprétation peut être différente. D'un point de vue légal en particulier, le bâtiment reste suffisamment lui-même pour que le droit de garder l'ancien coefficient d'occupation des sols lui soit accordé, au lieu de devoir adopter un nouveau coefficient plus petit. En revanche d'un point de vue fonctionnel, il s'agira d'un nouveau bâtiment.

Ainsi, outre le fait que, pour proposer aux utilisateurs une typologie enrichie des évolutions, le producteur de données doive développer une solution technique de suivi d'évolutions sophistiquée, il est délicat pour lui d'établir les spécifications nécessaires à l'interprétation des modifications. Ceci est particulièrement vrai lorsque les données produites sont génériques, destinées à un grand nombre d'utilisateurs et d'applications. Une interprétation pré-suppose d'une part d'être capable de statuer sur la continuité de l'objet géographique tel qu'observé par le thématicien (il faut donc dans certains cas être très proche de l'utilisateur final) et d'autre part d'être capable de suivre cet *objet* au travers du filtre de la formalisation technique (de l'*objet géographique* à l''objet informatique'). Une interprétation générique des données d'évolutions par le producteur non-utilisateur peut dès lors dans certains cas non pas réduire le « gap », mais au contraire l'accentuer en rigidifiant le cadre d'interprétation, rendant plus délicate pour l'utilisateur l'interprétation automatique des changements en terme d'évolutions.

En conclusion, les solutions techniques pour l'utilisateur consistent :

– soit à se « lancer » dans la production initiale des données qu'il souhaite observer. Cela lui permet de spécifier le modèle temporel qui lui convient, mais demeure une solution longue, coûteuse et en partie redondante par rapport aux bases de données déjà existantes,

– soit à développer les outils d'appariement et de calculs différentiels adaptés à son besoin, nécessaires à l'extraction des mutations à partir de

données existantes. Cela demande des compétences spécifiques en programmation et en logiciels SIG,

– soit à profiter de développements effectués par le producteur de données qui lui permettent d'accéder directement à des données d'évolutions. Cependant, il doit être conscient que les informations sur les évolutions fournies par ces données d'évolutions peuvent lui apparaître déformées par rapport à l'observation libre qu'il aurait pu en faire directement sur les objets géographiques et donc mal adaptées à l'observation du phénomène qu'il souhaite étudier. En effet, les données d'évolutions décrivent des 'objets informatiques' (objets observés par le producteur) et non les *objets* géographiques et par ailleurs, pour les interpréter, le producteur s'appuie sur des spécifications qui agissent comme un filtre sur les informations initiales.

B.3.1.2.3. Une approche méthodologique à l'entre-deux

Dans cette thèse, nous avons posé *l'entre-deux*, pour proposer une solution alternative au problème du suivi d'évolutions dans un SIG. Cette solution n'implique pas de produire de nouvelles bases de données mais propose d'exploiter les bases existantes. Centrée sur les informations géographiques considérées comme médiateur entre les données qui les représentent et les objets qu'elles modélisent, la solution utilise une méthode qui permet de définir des changements plus subtils et plus thématiques que ceux rigoureusement saisissables sur des données, tout en limitant la part de pré-interprétation. Dans ces travaux, nous suggérons de laisser l'interprétation des évolutions au soin des thématiciens et l'extraction des changements de données aux techniciens. Il peut arriver dans certaines structures que les deux rôles soient tenus par un même groupe de personnes, mais cette situation n'étant pas générale nous préférons distinguer les deux rôles. Avec la méthode d'observation multi-niveaux, nous proposons de construire des informations sous forme d'indicateurs qui soient faciles de suivre au cours du temps. Notre travail s'arrête avant l'interprétation des valeurs prises par ces indicateurs (par exemple, interprétation de valeurs limites) et de leurs « mouvements » (croissance, stagnation, régularité, inflexion…). Nous avons choisi d'appuyer la méthode en standard sur une configuration simple (le modèle par *snapshots*), l'exploitation de modèles et d'outils d'analyse de la dimension temporelle plus élaborés devenant alors un mode d'enrichissement optionnel.

B.3.2. Suivi de phénomènes géographiques dans un SIG avec la méthode d'observation multi-niveaux

B.3.2.1. Exploitation standard de la méthode

Pour observer l'évolution d'un objet géographique, il importe de disposer d'informations détaillées et de pouvoir les suivre. Dans le chapitre B.2, nous avons vu que l'utilisateur de SIG est souvent confronté au manque de données et donc au manque d'informations sur l'objet qu'il souhaite étudier. Nous avons expliqué comment en produire avec la *méthode d'observation multi-niveaux par emprise*. Dans ce qui suit nous allons montrer que les informations ainsi produites peuvent être suivies facilement et ce, même si le modèle temporel et les moyens d'analyses utilisés sont sommaires, donnant ainsi au thématicien la possibilité de suivre simplement des évolutions sur les objets qui l'intéressent. En standard, nous nous plaçons dans le cas de modèles par *snapshots* avec estampillage

6. Le modèle retenu « par snapshots » correspond à un cas simple d'un point de vue technique et assez courant dans la pratique. Par ailleurs, l'option « avec estampillage » qui consiste à affecter la date de la version de la base aux données qu'elle contient est une opération technique élémentaire qui apporte un vrai enrichissement pour un moindre coût (cf. A.3.2.3).

des données[6], sans outils d'appariement. Nous verrons alors que même dans ces conditions, la méthode permet de :

– disposer de données caractéristiques adaptées à un suivi d'évolutions (B.3.2.1.1) ;

– pouvoir suivre les informations dans le temps sans recourir à des modèles et outils sophistiqués (B.3.2.1.2) ;

– et éventuellement de pouvoir adopter des caractéristiques d'observation (*pas d'observation, date d'origine* de l'observation) de façon plus souple (B.3.2.1.3).

B.3.2.1.1. Des informations adaptées au suivi d'évolutions

La *méthode d'observation multi-niveaux par emprise* permet de construire des informations (au sens informatique, c'est à dire comme résultat de traitement de données) qui possèdent plusieurs qualités pour servir au suivi d'évolution :

- les informations sont facilement observables dans un SIG ;
- les informations viennent enrichir la description de l'objet étudié pour une observation plus étoffée ;
- les informations donnent des informations de détails sur l'objet étudié pour permettre de suivre des évolutions et pas seulement pour observer des changements ;
- les informations sont fournies sous forme d'indicateurs laissés à l'interprétation de l'utilisateur.

Des informations facilement observables dans un SIG

Précédemment nous avons montré que la méthode proposée permet de créer des informations supplémentaires. Par construction, ces informations viennent décrire l'objet au *niveau adéquat d'observation* et la *gamme d'échelle* d'utilisation des données produites est ainsi adaptée à l'échelle d'observation souhaitée par l'utilisateur. Par ailleurs, les informations produites avec la méthode proposée sont données sous la forme d'attributs d'un *objet géographique par nature* représentant l'entité géographique sous forme d'un objet au niveau méso. Ces informations sont donc directement accessibles. Ces apports de la méthode d'observation multi-niveaux (*gamme d'échelle* adaptée pour une observation au *niveau adéquat*, à partir d'*informations explicites* directement *accessibles* par requêtes standard élémentaires) facilitent de fait l'observation de phénomène dans un SIG.

Des informations qui viennent enrichir les observations

La méthode que nous présentons, en s'appuyant sur une *agrégation des informations par emprise* assure une double compatibilité :

. la compatibilité entre les informations sur l'objet étudié pouvant être disponibles par ailleurs au niveau méso, et les informations méso-structurelles construites à partir de l'agrégation d'informations micro,

. la compatibilité, entre elles, des informations méso-structurelles issues de données micro variées.

La méthode permet ainsi d'intégrer dans une même observation des informations existantes et les informations construites, des informations nouvelles issues de bases de données différentes portant sur des thèmes variés. Elle peut servir à introduire une plus grande diversité d'indicateurs pour une observation plus riche. Elle peut ainsi servir de support à une appréhension pluridisciplinaire du même objet.

Des informations qui permettent de saisir des « changements insensibles »

Dans le contexte d'une observation cinématique, les informations produites par la méthode d'observation donnent une base empirique au concept de « l'objet suffisamment lui-même ». En effet, l'agrégation d'objets micro ne nous sert pas seulement à fournir plus d'informations à l'observation en enrichissant la description de l'objet méso. Elle nous sert surtout à affiner cette description pour appréhender l'objet à une échelle plus fine. A cette échelle, les changements observés peuvent être plus subtils. Ainsi, en donnant à voir des changements (micro), la méthode fournit à l'utilisateur des éléments pour mieux interpréter les changements méso et pour distinguer des formes d'évolutions plus nombreuses (changements graduels/changements brusques).

Des informations qui s'expriment sous forme d'indicateurs à interpréter

Les informations construites, en tant que descripteurs d'agrégations d'objets micro (ou d'agrégations de caractéristiques d'objets micro) sont obtenues essentiellement à partir de calculs mathématiques sur des données micro. Les valeurs qui en résultent sont donc des valeurs brutes qui peuvent être fournies telles quelles à l'utilisateur pour limiter les effets d'une pré-interprétation.

B.3.2.1.2. Un lien temporel entre les données

Condition initiale d'application de la méthode multi-niveaux pour une observation cinématique

Le chapitre B.2 a développé la mise en œuvre de la méthode pour observer statiquement un objet. En supposant maintenant :

– que *l'emprise au sol de cet objet puisse être considérée comme stable sur la période d'observation* [7]
– que *le support d'agrégation 'portion de territoire' soit le même pour toutes les bases de données micro,*

nous allons voir ici que la méthode d'observation multi-niveaux permet de suivre facilement les indicateurs produits, sans impliquer l'utilisation d'un modèle spécifique de gestion des mises à jour ni le développement d'outils d'appariement.

Rappel des principes de l'observation multi-niveaux statique

Le principe de la *méthode d'observation multi-niveaux* consiste à :

– décrire la portion de territoire couverte par l'objet géographique étudié, à l'aide de caractéristiques pertinentes pour le phénomène suivi,

– puis à transférer à l'objet géographique les caractéristiques du territoire qu'il occupe, par identification de l'objet géographique à son emprise.

La mise en œuvre de la méthode exploite des données sur des objets micro pour caractériser la 'portion de territoire' associée à l'objet méso géographique : Elle utilise une *agrégation par emprise* pour sélectionner l'ensemble des objets micro appartenant à cette 'portion de territoire' et calcule ensuite différentes caractéristiques relatives à cet ensemble. Ces caractéristiques considérées comme des attributs de la portion de territoire sont ensuite intégrées aux attributs de l'objet méso (par exemple, on dira que la zone pavillonnaire possède *n* bâtiments et que la surface moyenne du bâti de cette zone est de *x* m^2).

7. On peut argumenter à propos de la contrainte fixée par l'hypothèse. Poser une emprise au sol constante pour l'objet suivi au niveau méso peut en effet paraître restrictif. L'application de la méthode implique soit de se trouver dans cette situation soit de s'y ramener.
Dans les faits, cette hypothèse s'avère moins contraignante qu'il n'y paraît. En effet, dans certains cas elle correspond aux conditions d'observations définies par l'utilisateur lui-même, lorsqu'il observe les évolutions d'une portion de territoire spécifique sous l'effet d'un certain phénomène (ex. : l'évolution du foncier dans une zone suite à l'installation d'une ligne de tramway ou le développement de l'activité commerciale suite à l'ouverture d'une station de métro) ; La zone d'étude reste constante et l'utilisateur regarde ce qui s'y produit.
Dans d'autres cas, en particulier pour des phénomènes urbains, les objets méso suivis (rue, îlot, quartier mais aussi commune, région, etc.) sont des objets relativement stables dans le temps (car ils s'appuient sur des infrastructures (routes, bâti) ou sur des limites elles-mêmes pérennes) permettant ainsi d'appliquer la méthode…
Toutefois, pour certaines applications, il est nécessaire d'effectuer un travail d'analyse et de modélisation préalable pour s'adapter à la contrainte de départ. Même si la difficulté n'est pas technique, elle peut être réelle pour le thématicien ; Il s'agit d'identifier un *bon* objet méso d'étude. Ce travail de modélisation en informations géographiques peut pour les cas les plus délicats nécessiter l'aide d'un géomaticien. Le cas de l'ambiance urbaine où il a fallu trouver un objet méso qui ait un sens thématique pour suivre le phénomène d'*ambiance urbaine* en est une illustration.

D'autres solutions existent pour s'adapter aux conditions d'applicabilité de la méthode, en particulier le découpage en objets méso intermédiaires, présenté comme élément d'enrichissement de la méthode d'observation statique permet d'envisager le traitement des modifications de type extension / rétraction... Dans l'Annexe 5, nous prolongeons cette discussion sur l'hypothèse de l'emprise constante.

Observation cinématique pour le suivi d'évolutions

Pour effectuer une observation cinématique discrète d'un *phénomène*, il faut disposer d'un ensemble d'états et d'un lien entre eux permettant de les ordonner en une suite temporelle (fig. B.3.7a).
Pour observer cinématiquement un *objet dans un SIG*, il faut avoir des représentations d'états différents sur cet objet et posséder les moyens d'établir une relation temporelle entre ces états (fig.B.3.7b).
Pour observer des évolutions *détaillées* de l'objet méso géographique, il faut disposer de descriptions d'états détaillées et de moyens pour établir un lien entre les valeurs des caractéristiques utilisées pour décrire ces états (fig. B.3.7c) ; Autrement dit pour chaque caractéristique, il doit exister un lien entre les différents états de cette caractéristique pour permettre un suivi détaillé *automatique* d'un état à l'autre.

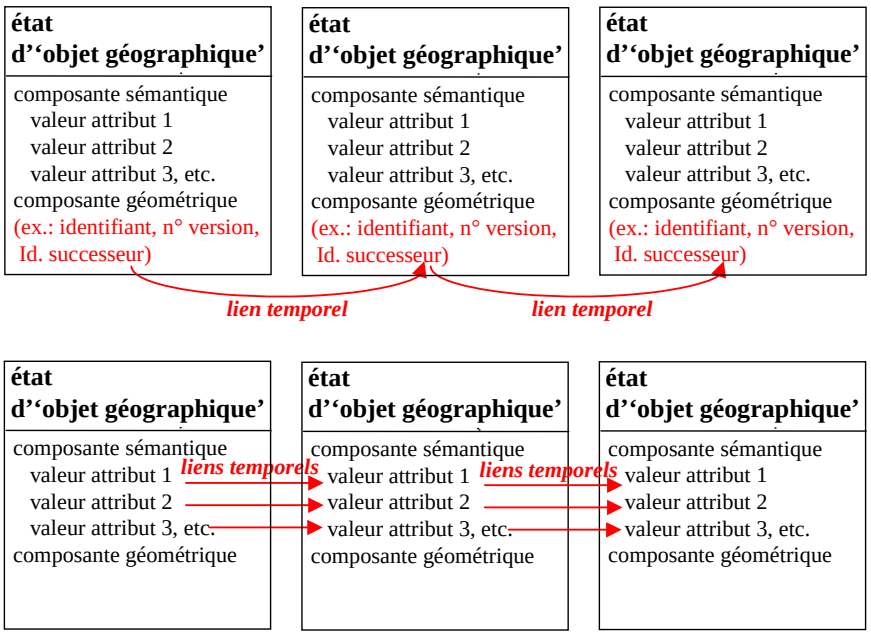


Figure B.3.8. *Le suivi d'évolution dans un SIG*

Le suivi d'évolutions dans un SIG nécessite un lien automatisable entre les états (d'objets ou de caractéristiques)

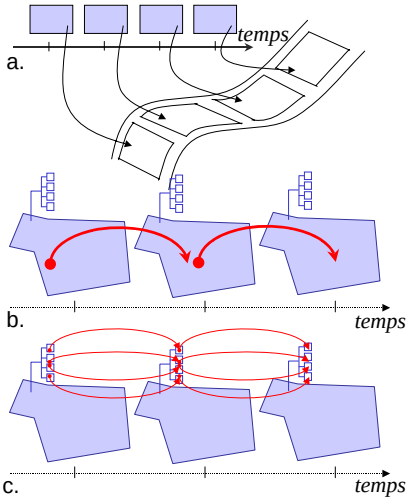


Figure B.3.7. *Observations cinématiques*

a. suivi cinématique d'un phénomène
b. suivi dans un SIG des états d'un objet
c. suivi dans un SIG des attributs d'un objet

Dans les modèles de gestion des mises à jour, lorsqu'il existe un lien temporel, il porte en général sur des *états d'objets* et non sur des *états d'attributs d'objet*[8] (cf. figure B.3.8). De fait, cela contraint l'utilisateur qui veut faire du suivi d'évolution à analyser « manuellement » la description des états d'objet pour identifier quelles sont parmi ces caractéristiques celles qui ont changé. Les outils de calculs de différentiels couplés à des outils d'appariement facilitent l'extraction automatique des changements, mais la complexité de mise en œuvre et leurs limitations pour traiter certains cas d'appariement réduisent aujourd'hui leur exploitation.

Dans ce qui suit nous montrons que la méthode d'observation multi-niveaux par emprise permet de suivre des évolutions à partir du modèle sommaire dit par *snapshots*. La mise en œuvre d'un suivi d'évolution avec la méthode multi-niveaux se contente de reproduire une observation statique dans le temps, sans se préoccuper d'établir de lien entre les états des attributs observés ; le lien entre les informations d'état résulte directement de la méthode de production des indicateurs.

Mise en œuvre de la méthode multi-niveaux pour un suivi d'évolutions

Pour construire les descriptions d'états nécessaires au suivi de l'objet méso géographique étudié, nous utilisons la méthode multi-niveaux en itérant les

8. Par exemple :
- Le modèle par historique propose un identifiant du successeur ;
- Le modèle par versionnement utilise un ‘numéro de version’ d'objet pour passer d'un ‘état d'objet’ à celui qui lui succède.
Seul le modèle par « versionnement sans description totale » (qui n'enregistre que la valeur des caractéristiques qui ont été modifiées entre deux états, cf. § A.3.2.3.2) permet de déduire par complémentarité que les valeurs qui n'ont pas été mises à jour sont restées dans un état ‘identique’ et pour celles qui ont été modifiées de retrouver par requête leur valeur dans l'état précédent.

étapes d'une observation statique. Dans ce qui suit nous avons simplifié les notations pour faciliter la compréhension du principe[9].

Dans un premier temps :

- Nous répercutons sur les données l'hypothèse de stabilité de l'emprise au sol de l'objet géographique : la 'portion de territoire' associée à l''objet méso géographique' est ainsi posée comme constante sur le temps de l'observation (fig. B.3.9.I).

- Nous introduisons ensuite la base des données micro. Ayant supposé qu'elle exploite un modèle par *snapshots*, nous considérons les versions v_t de cette base aux différentes dates t (fig. B.3.9.II). Pour chaque version v_t, nous déterminons par agrégation l'ensemble des objets micro issus de la version v_t inclus dans la 'portion de territoire associée à l'objet méso géographique' à la date t. Nous le noterons $\{o_i\}_t$ (fig. B.3.9.III).

- Enfin, nous définissons parmi les caractéristiques relatives aux ensembles d'objets micro celles que nous souhaitons établir (f_k) (ex. : cardinal de l'ensemble, surface moyenne des objets, etc.). Ces caractéristiques obtenues par calculs (pouvant combiner des fonctions statistiques, géométriques, topologiques) se présentent comme des formules mathématiques $f(o_i)$ fonctions des objets micro $\{o_i\}$ agrégés.

Par exemple, $f(\{o_i\}) = \dfrac{\Sigma \textit{ Surface des } o_i}{\textit{Nombre de bâtiments dans la portion de territoire}}$ où i décrit l'ensemble des bâtiments inclus dans la portion de territoire

peut être utilisée dans le cadre d'une observation cinématique d'une zone pavillonnaire, pour obtenir la valeur de la caractéristique 'surface moyenne des bâtiments de la zone'.

Dans un deuxième temps :

- Pour chaque date t, nous calculons la valeur $f_k(o_i)_t$ de chaque caractéristique f_k pour l'ensemble des objets micro $\{o_i\}_t$ résultant de l'agrégation par emprise à la date t. Les f_k étant considérées comme des attributs de la 'portion de territoire associée à l'objet méso géographique', les valeurs $f_k(o_i)_t$ correspondent alors aux états de ces attributs aux dates t (fig. B.3.9.IV).

- Enfin, après identification de l'objet méso géographique à la portion de territoire qu'il occupe (fig. B.3.9.V), les caractéristiques sont intégrées aux attributs de l''objet méso géographique' et leur valeur aux différentes dates t, sont alors considérées comme les 'états de ces attributs' pour les différents 'états de l'objet méso géographique' (fig. B.3.9.VI).

Au final, nous obtenons différents 'états temporels' de l''objet géographique' étudié, où chaque état est défini par l'ensemble des valeurs prises par les indicateurs à la date correspondante.

9. Dans l'Annexe 6, nous présentons cette étape en utilisant une formulation mathématique détaillée plus rigoureuse, qui intègre sous forme d'indices différents les variations : du temps (t_p), de caractéristique considérée (f_k), des objets micros agrégés (o_j), ainsi que de l'objet méso observé (O_j).

I. La 'portion de territoire' associée à l'objet est posée comme constante

II. Soient les versions v_t de la base de données micro,…

III…. les données sont agrégées en fonction de l'emprise de l''objet méso',

IV. On calcule la valeur des caractéristiques f_k pour chaque ensemble agrégé d'entités,

V. Après identification à la portion de territoire commune, …

VI. …on transfert les caractéristiques calculées sur les agrégations à l''objet méso géographiques' observé dans le SIG.

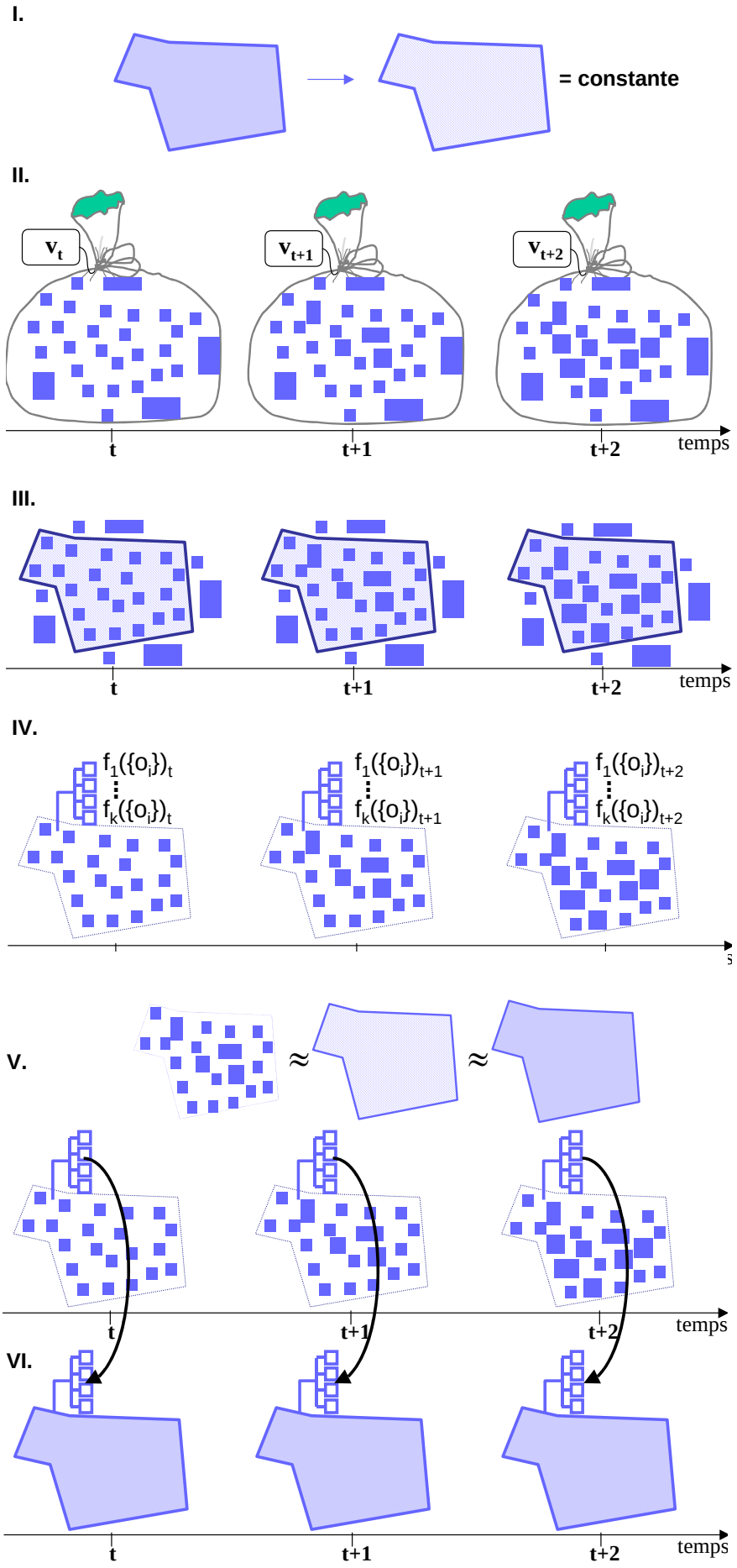


Figure B.3.9. *Mise en œuvre de la méthode multi-niveaux pour une observation cinématique*

Nous allons maintenant constater que sans intervention supplémentaire, la série des observations statiques obtenue par cette méthode permet d'effectuer une observation détaillée des évolutions de l'objet suivi dans un SIG. Il existe un lien simple à exploiter et automatisable entre les différents 'états' ('états d'objet' *et* 'états de caractéristiques') qui permet ainsi une observation cinématique fine dans un SIG.

Un lien implicite entre les états, simple à exploiter

L'existence d'un lien entre les *états de l'objet* et surtout entre les *états de ses caractéristiques,* qui soit exploitable dans les SIG standards, résulte de trois faits conjugués :

– Les 'états d'objet' et les 'états de caractéristique' correspondent à des descriptions en fonction du temps, respectivement du même 'objet' et du même 'indicateur' sur le même 'objet'. Il y a donc *continuité* de la « chose observée ». Cela a donc du sens de comparer ces descriptions et de tenter de les suivre pour observer leur évolution[10].

– La gestion des 'états de caractéristiques' dans un SIG est simplifiée par la formulation mathématique des indicateurs servant à leur calcul ;

– La forme cinématique de ces indicateurs (donnés comme une fonction du temps) établit implicitement le lien temporel nécessaire au suivi des différents 'états' de ces indicateurs.

L'hypothèse d'emprise constante entraîne la comparabilité des indicateurs dans le temps

La méthode d'observation multi-niveaux utilise l'emprise de l'objet méso observé (ou, quand elle n'est pas disponible, une emprise de substitution correspondant à l'objet méso-structurel cf. B.2 p. 134-143) pour définir des ensembles d''objets micro'. Elle caractérise ensuite ces ensembles pour obtenir des informations sur l'objet (méso) à un niveau de description (micro) plus fin. Avec l'hypothèse de constance de l'emprise, nous imposons l'identité de l'emprise d'agrégation. Les valeurs prises par les caractéristiques deviennent alors comparables dans le temps : elles représentent les 'états' d'un même indicateur donné par la fonction f_k sur le même objet 'portion de territoire', les 'objets micros' qui participent au calcul de cette fonction sont alors considérés comme des paramètres. Intégrées aux attributs de l''objet observé', ces valeurs représentent alors les 'états' d'une même caractéristique et leur suivi peut donc participer au suivi des évolutions de cet 'objet'.

La forme des indicateurs simplifie la gestion temporelle de l'information

Il s'agit maintenant de pouvoir effectuer ce suivi simplement. Bien que la gestion des informations temporelles soit généralement délicate dans les SIG, elle se fait ici sans impliquer de modèle particulier. En effet, la série temporelle des valeurs prises par chaque caractéristique résulte de calculs successifs ; Nous avons vu qu'elle pouvait s'exprimer sous la forme de $f_k(o_i)t$.

Nous conseillons ici d'intégrer ces valeurs sous forme de valeurs d'attributs dans la série des représentations d'états de l'objet : le fait que ces valeurs soient stockées « en dur » (c.a.d qu'elles fassent partie des valeurs d'attributs contenues explicitement dans la composante sémantique) facilite leur exploitation dans le SIG (pour leur consultation, pour leurs représentations cartographiques, etc. cf. B.1, p. 110). Cependant, conceptuellement, chaque caractéristique peut être gérée par ailleurs, comme une formule mathématique renvoyant pour chaque état à la valeur des différents

10. La condition de *continuité de l'objet observé* est indispensable pour envisager un suivi d'évolutions. Il s'agit ici de valider la comparabilité des informations construites. En effet, pour pouvoir comparer des 'états d'objet' (et des 'états de caractéristique'), il ne suffit pas qu'ils possèdent le « même intitulé ». Il faut qu'ils représentent le même objet (ou caractéristique) et que les représentations soient cohérentes.

En effet, on peut se trouver avec des descriptions qui ne sont pas comparables parce que l'objet référent n'est pas le même :

- soit parce que la définition de l'entité est intrinsèquement floue et qu'en croyant saisir le même objet, on observe des objets différents, par exemple la *ville* (*« le terme de ville, à la différence de ceux d'agglomération ou de banlieue par exemple, n'a pas de définition précise »* [Nicot, 1996, p289]).

- soit parce que chaque utilisateur à une définition précise de l'objet mais celle-ci diffère d'un utilisateur à l'autre (ex. : la zone inondable dans la BD TOPO® est *« une zone de terrain naturel pouvant être recouverte d'eau »*. Comme le note les spécifications de la base : *« l'objet ZONE_INONDABLE ne correspond pas aux zones du territoire qui peuvent accidentellement être inondées lors de certaines crues, ni aux zones déclarées administrativement inondables donc non constructibles, mais aux zones marquées de façon permanente par le passage ou la présence prolongée de l'eau »* [IGN, 1999, p.121].

Il en résulte qu'un même attribut peut se trouver dans des états très différents sans qu'il soit question d'évolution.

Par ailleurs, on peut se trouver avec des descriptions d'objet cohérentes mais ne pas pouvoir comparer les valeurs de leurs caractéristiques. Ainsi pour l'objet *cours d'eau,* on peut utiliser la caractéristique *largeur* et la définir à l'aide d'indicateurs différents (ex. : en fonction d'une plages de valeurs {< 5 m ; 5 à 15 m ; > 15m} ou d'une autre {< 7 m ; 7 à 20 m ; 20m ; > 20 m}). Ce qui rend de fait difficile toute comparaison [Mézard, 1994].

paramètres[11]. Dans un premier temps, cette gestion simple, indépendante des SIG, libère l'utilisateur de la nécessité d'un modèle de gestion des mises à jour élaboré. Dans un second temps, elle évite d'avoir recours à des outils d'appariements et de calculs différentiels pour déterminer les changements entre deux 'états de l'objet' : Pour chaque couple d'états, ils se déduisent aisément du calcul de $f_k(o_i)t_{p+1}$ - $f_k(o_i)t_p$ pour l'ensemble des caractéristiques f_k. Sans difficulté technique supplémentaire nous pouvons également suivre individuellement les états de chacun des indicateurs f_k que nous avons construit à partir des séries $\{f_k(o_i)t_0 ; f_k(o_i)t_1\} ; \ldots ; f_k(o_i)t_p\}$; $f_k(o_i)t_{p+1}\ldots\}$. Nous disposons donc ainsi d'un lien entre les 'états des caractéristiques'.

Une formulation cinématique des caractéristiques observées

La méthode d'observation multi-niveaux exprime les caractéristiques observées sous forme d'une formule mathématique fonction du temps. Elle les présente ainsi directement sous une *forme cinématique* (cf. A.2, p. 54). Conceptuellement, il est alors possible d'effectuer une observation de ces indicateurs à toute date t ; il suffit pour cela de disposer pour cette date t des 'états des objets micros' utilisés comme paramètres dans les f_k… Il en résulte surtout la possibilité de déduire le lien entre les différents 'états' de ces indicateurs à partir du lien temporel entre les dates d'observation : Ayant un ensemble d'observation $f_k(o_i)t_p$ pour établir la série temporelle, il suffit de les ordonner en fonction des dates t_p.

Nous venons de voir comment la modélisation des indicateurs sous la forme cinématique d'une fonction du temps nous permet de disposer d'un lien temporel entre les données. Nous allons maintenant montrer que cette modélisation peut également permettre d'assouplir le pas d'observation.

B.3.2.1.3. Un pas d'observation plus souple

Contraintes sur le choix du pas d'observation des données

Le choix d'un pas d'observation doit être adapté au phénomène étudié. Toutefois pour des raisons techniques, il est généralement contraint par les bases de données exploitées ; Ainsi, l'observateur s'adapte souvent au pas des données qu'il utilise mais qu'il ne produit pas. Par exemple, pour observer un phénomène à partir d'une base de données s'appuyant sur le modèle par *snapshots*, l'utilisateur peut adopter au mieux un pas d'observation correspondant au pas de mise à jour de la base : En deçà, il n'y a aucune information nouvelle à observer[12].

L'utilisation de données variées dans une même application accentue encore le problème du choix du pas d'observation. En effet, l'intégration de bases de données diverses pose souvent des problèmes de cohérence : cohérence des spécifications renvoyant à l'intégration de *représentations multiples*, mais aussi cohérence temporelle des dates d'observation. En effet, deux bases de données construites indépendamment par des producteurs différents ont une forte probabilité d'avoir des dates de versions différentes[13]. Lorsque c'est le cas, l'utilisateur hésite souvent à les exploiter simultanément dans la même observation, craignant peut être que la double incompatibilité :

- celle des descriptions liées à une représentation multiple (en particulier celle des supports géométriques),
- celle des supports d'observations temporelles,

n'introduisent des erreurs trop importantes dans les observations.

11. Certains logiciels SIG offrent la possibilité de modéliser « dynamiquement » certains attributs : elle assure la continuelle actualité de l'attribut automatiquement, au prix cependant de l'« écrasement » systématique des anciennes valeurs. Cela consiste à modéliser un attribut sous forme de fonction, à utiliser des formules plutôt que des valeurs constantes. Typiquement, les indicateurs que nous construisons seraient susceptibles d'une telle modélisation, sous réserve d'une observation uniquement statique ou sinon, d'un modèle d'archivage spécifique.

12. La mise à jour en continu, qui consiste à mettre en place des structures de veille pour identifier au plus tôt les changements intervenus sur le terrain et à les entrer « en temps réel » dans les bases de données convient particulièrement bien aux utilisateurs. En effet, elle assure une plus grande homogénéité de l'actualité de la base au cours du temps et elle permet à l'utilisateur d'adopter un pas d'observation adapté à ses besoins. Le pas de collecte des informations pouvant être considéré comme nul, l'utilisateur n'est pas contraint d'adopter un pas coïncident avec un pas d'observation qui ne lui conviendrait pas.
En fait, l'introduction des changements en « temps réel » devrait permettre une observation en temps réel. Toutefois, la granularité temporelle des opérations d'acquisition des informations (c'est à dire le temps nécessaire à la collecte et à l'intégration de ces informations) introduit un délai dans l'observation et donc un pas d'observation minimal.

13. Pour l'application *zone pavillonnaire*, nous avons eu de la chance : le MOS et la BD TOPO® avaient des temporalités à peu près compatibles.

Un pas d'observation des informations plus souple

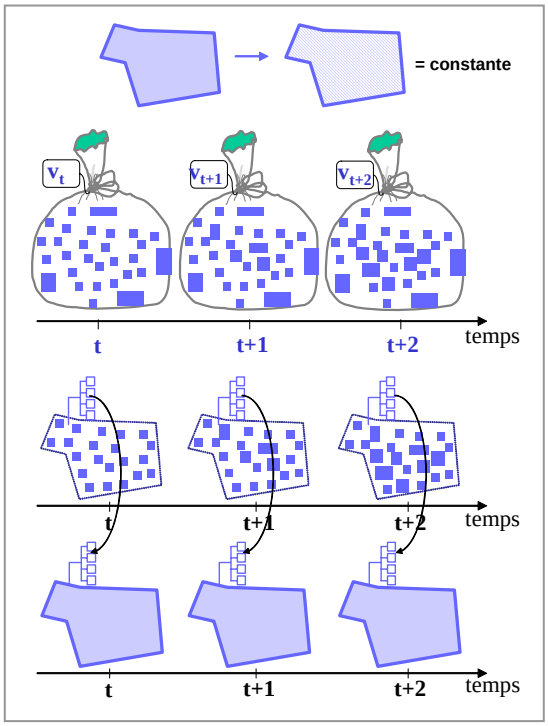
En abordant le problème en terme *d'informations* plutôt qu'en terme de *données*, il nous semble pouvoir être appréhendé autrement. Tout d'abord nous remarquons que, dans la pratique, une version de base de données est considérée comme « à jour », jusqu'à l'arrivée de la nouvelle version ; Les utilisateurs exploitent en effet la dernière version de la base de données, (c'est à dire, ils effectuent leurs analyses, leurs observations sur les données les plus récentes) comme s'il s'agissait de représentations actuelles[14]. Cette pratique revient à considérer que sur l'intervalle de temps $[t_p ; t_{p+1}[$, l'état des objets (en tant qu'information) reste constant. Elle peut aussi être vue comme une *extrapolation de l'état des objets de la base et des états de leurs caractéristiques par la fonction identité*, jusqu'à l'obtention de données nouvelles.

Conceptuellement, l'extrapolation (ou l'interpolation) de bases de données géographiques vecteurs semble compliquée : *Comment extrapoler des maisons, un réseau routier, des cours d'eau, des communes ?* En revanche, l'extrapolation de valeurs d'attributs comme celles des indicateurs définis par la méthode multi-niveaux est plus facilement concevable. Nous avons vu précédemment que la valeur de ces caractéristiques pouvait être calculée à tout instant t, à partir d'une version de la base micro. Nous proposons ici d'envisager le calcul d'une estimation de ces valeurs lorsque aucune version de la base n'est disponible.

Ainsi, si nous considérons les versions $\{v_{tp}\}$ et $\{w_{tp'}\}$ de deux bases de données micro correspondant à des dates différentes $\{t_p\}$, $\{t_{p'}\}$, alors nous pouvons construire des caractéristiques sur l'objet méso aux différentes dates $\{t_p\}$ à partir des $\{v_{tp}\}$ en appliquant la méthode multi-niveaux (de même aux dates $\{t_{p'}\}$ à partir des $\{w_{tp'}\}$) (cf. figure B.3.10.I). Comme évoqué précédemment, on peut travailler entre deux mises à jour « comme si » les données restaient constantes. Après avoir attribué à l'objet méso l'ensemble des caractéristiques obtenues à partir des deux bases, on obtiendra alors une représentation unique de l'objet méso géographique. Cette représentation enrichie des deux ensembles de caractéristiques proposera des représentations d'états différents à des fréquences plus élevées (cf. figure B.3.10.I). Ce qui par définition autorise un pas d'observation plus fin (cf. bas de la figure B.3.10.II). Toutefois, l'objet continuera à évoluer sur les deux ensembles de caractéristiques de façon indépendante, selon deux rythmes différents correspondant aux rythmes de mise à jour de chacune des bases micro, donnant dès lors à voir une représentation cinématique de l'objet méso un peu « schizoïde ». Par contre, si l'utilisateur est capable « d'interpoler (et d'extrapoler) les 'états de l'objet' » entre deux versions de base de données micro, (ce qui dans les faits se traduira par l'interpolation des 'états des caractéristiques'), alors pour toutes les dates t de l'intervalle $[t_p ; t_{p+1}[$, il pourra calculer une interpolation de la valeur des indicateurs obtenus à partir des versions v_{tp} et v_{tp+1}, permettant ainsi d'obtenir sur l'ensemble des dates une représentation de l''objet' plus unifiée.

14. Ainsi, pour l'observation statique, nous avons travaillé sans prendre en compte la dimension temporelle, sans envisager de décalage temporel entre la date de l'observation statique et les dates des versions des bases. Implicitement, nous utilisons la « meilleure version de la base» au sens de celle qui donne la meilleure représentation en fonction de la date. Nous calculons les caractéristiques et nous les intégrons ensuite pour une même observation. Nous « acceptons » d'utiliser des versions des bases micros, dont la date de mise à jour associée ne correspond probablement pas exactement à la date de l'observation statique, en considérant que l'état des objets est le même que dans la version de la base la plus proche Ce qui équivaut à utiliser une fonction *identité* comme fonction d'interpolation (ou d'extrapolation) des états de ces bases autour de la date d'observation.

Principe général de la méthode d'observation multi-niveaux



« L'emprise de l'objet méso observé étant posée constante, une base de données micro étant disponible en plusieurs versions, on calcule des indicateurs pour caractériser les agrégations d'objet s'appuyant sur l'emprise considérée et on les affecte alors à l'objet méso étudié. »

Application du principe général de la méthode d'observation multi-niveaux, en parallèle à deux bases de données micro ayant des dates de versions

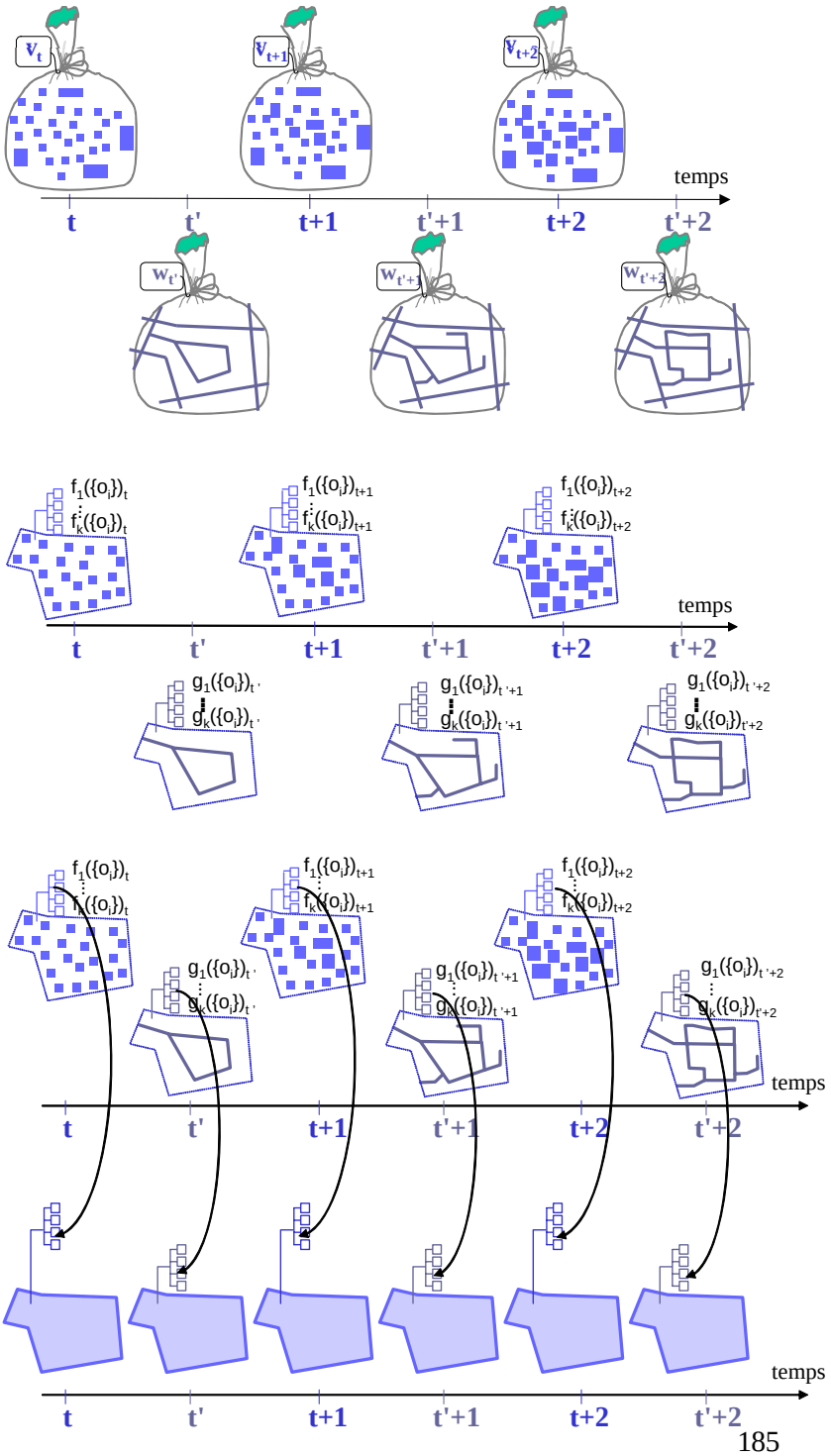


Figure B.3.10. I. *Intégration basique de données ayant des dates de versions différentes*

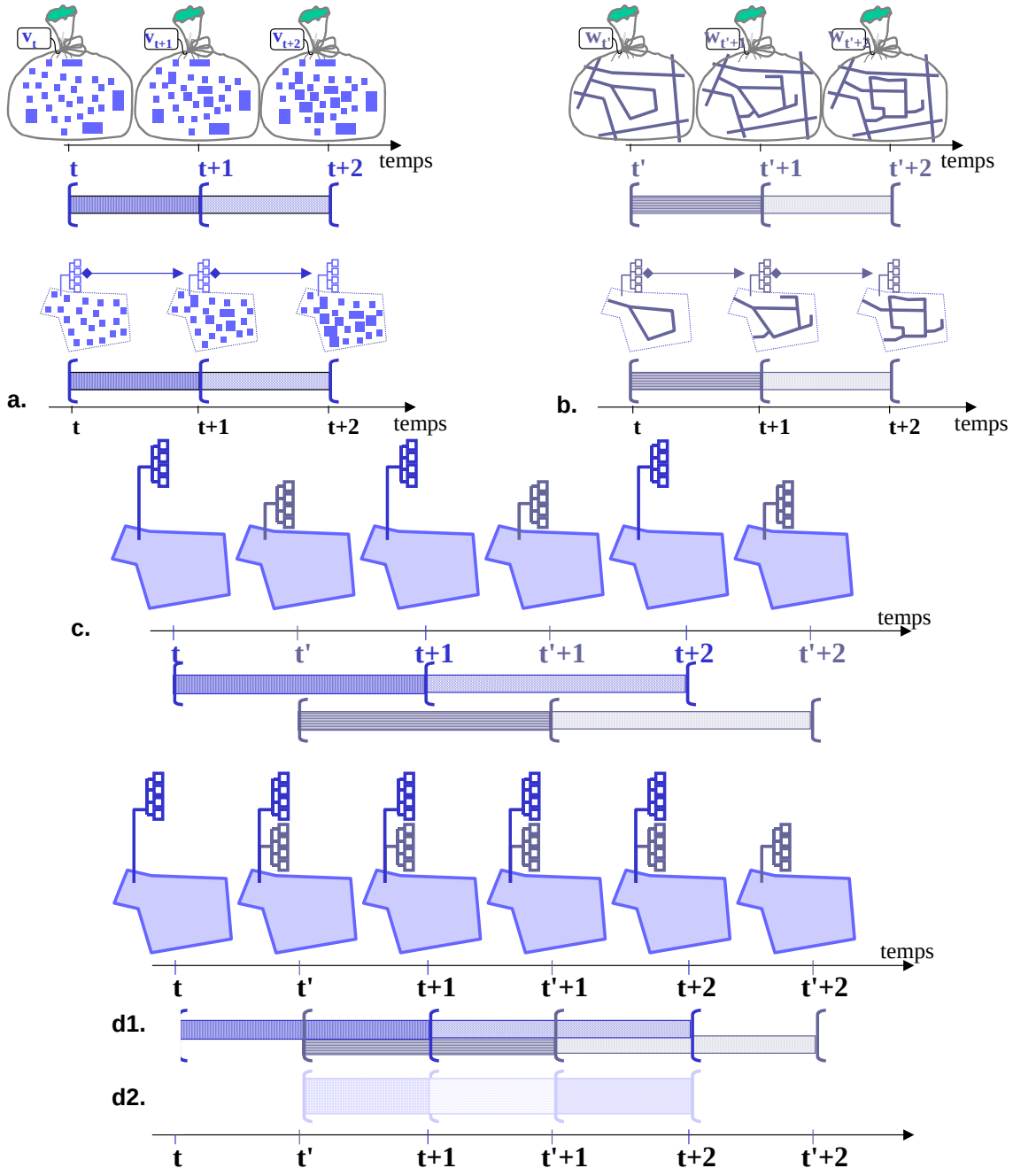


Figure B.3.10. II. *Intégration complète : utilisation de la fonction identité*

Considérons deux bases de données ayant des dates de versions différentes que l'on veut observer avec la méthode d'observation multi-niveaux et supposons qu'entre deux dates de mise à jour, les données de chaque base puissent être considérées comme constantes (a) et b).

Alors, pour une base de données fixée, on peut affecter à l'objet méso observé les valeurs d'indicateurs calculées aux dates correspondant aux mises à jour (c), mais on peut également compléter d'autres descriptions à d'autres dates par interpolation (d1). On obtient ainsi une description complète de l'objet à un « rythme de temps » plus réduit (d2).

Conceptuellement, la méthode multi-niveaux permet de concevoir des caractéristiques à observer :

- à n'importe quelle date
- à partir de bases de données micro exploitant un modèle par *snapshots* et de fonctions d'interpolation (et d'extrapolation),
- avec un lien temporel implicite entre les états de ces caractéristiques qui permette de les suivre pour une observation cinématique.

Dans la pratique, elle facilite l'intégration dans une même observation d'informations d'origines variées. Elle peut également permettre une plus grande souplesse d'observation à la condition de considérer la « fonction identité » comme une fonction d'interpolation valide ou de disposer d'autres fonctions d'estimations plus pertinentes.

Exploitation pour des observations rétrospectives

La méthode d'observation permet également d'envisager des observations rétrospectives, à la condition que l'emprise de l'objet observé reste constante. Pour cela on pourra :

– ou bien exploiter des données micro décrivant des époques antérieures à la date d'origine des observations et calculer des valeurs vraies d'indicateurs par application de la méthode multi-niveaux sur ces données supplémentaires ;

– ou bien calculer des estimations des caractéristiques observées par extrapolation des valeurs déjà fournies par la méthode.

Ce prolongement de la méthode d'observation, pour une observation rétrospective des évolutions peut intéresser les utilisateurs en particulier pour tester des modèles d'évolutions.

B.3.2.2. Amélioration de la méthode d'observation multi-niveaux cinématique

B.3.2.2.1. Analyses statistiques des indicateurs

Partant du besoin d'informations pour observer un phénomène, nous avons élaboré une *méthode générique* pour en produire. L'observation avec un SIG d'un phénomène géographique particulier implique ensuite d'élaborer une *méthode spécifique*. Il s'agit ici principalement de travailler avec l'utilisateur pour :

- définir le niveau adéquat d'observation du phénomène et les objets méso géographiques à partir desquels s'effectuera le suivi de ce phénomène,
- recenser les objets micro pouvant servir à caractériser le phénomène,
- déterminer un ensemble d'indicateurs pouvant servir à caractériser le phénomène qui soient déductibles des agrégations des données micros ;
- et parallèlement, d'étudier les données disponibles, leurs qualités et leurs modèles, les ressources et compétences mobilisables pour configurer les éléments principaux de la *méthode spécifique* (en particulier pour construire l'objet méso structurel utilisé comme emprise d'agrégation et pour calculer des indicateurs intéressants).

La mise en œuvre de *la méthode spécifique* permet alors à l'utilisateur de disposer d'un ensemble d'indicateurs. En statistiques, ces indicateurs se présentent classiquement sous la forme d'une série de matrices décrivant les valeurs prises par différentes caractéristiques sur différents individus au cours du temps. L'utilisateur pourra donc appliquer des méthodes d'analyses statistiques pour en déduire de nouvelles informations. Cependant, certaines analyses[15] ne sont applicables que si l'on dispose d'un

15. …comme celles qui, pour étudier l'évolution des caractéristiques en tenant compte de l'influence des états antérieurs, utilisent une modélisation probabiliste des objets en considérant chaque caractéristique comme un processus.

nombre suffisant d'observations. Ce nombre (de l'ordre de 100 mesures) ne peut résulter que d'une instrumentalisation et d'une automatisation des mesures rarement opérationnelles dans le domaine des phénomènes urbains {Bonin, 2006}. En revanche, d'autres analyses purement descriptives sur les caractéristiques (analyses en composantes principales sur les observations statiques) permettent d'étudier la dépendance des caractéristiques. Pour introduire le temps, on peut également calculer les variations entre chaque valeur successive des caractéristiques et étudier ces nouvelles variables à l'aide d'indicateurs classiques (moyenne, variabilité, fréquence) et de mesures d'indépendances entre ces variables... L'analyse statistique avancée des indicateurs fait appel à une expertise indépendante des SIG. Par ailleurs, elle ne peut intervenir qu'une fois la *méthode générique* et ensuite la *méthode spécifique* mises en œuvre, puisqu'elle opère sur les valeurs calculées des indicateurs. Elle ne fait donc pas partie de la *méthode générique*, elle lui fait suite. Toutefois, l'analyse statistique permet de déduire des caractéristiques produites de nouvelles informations qui peuvent à leur tour être observées et devenir ainsi des indicateurs pour le suivi des évolutions. A ce titre, elle constitue pour nous un mode d'enrichissement possible de la méthode.

B.3.2.2.2. Enrichissements des informations temporelles observées

Dans le paragraphe B.2.6 nous présentons différents moyens d'enrichir les informations descriptives pour une observation statique plus détaillée. Ces moyens profitent évidemment à une observation cinématique, puisqu'ils donnent :
- plus de détails à observer,
- plus de relations potentielles à identifier entre ces informations et leurs évolutions.

Il existe par ailleurs trois moyens spécifiques d'enrichir la méthode pour une observation cinématique :

– l'intégration d'une fonction d'interpolation entre deux valeurs calculées ;

– l'intégration de données supplémentaires utilisant un support d'agrégation différent ;

– l'intégration de moyens informatiques plus sophistiqués (modèles de gestion des mise à jour et/ou outils de calculs différentiels).

Ces enrichissements seront présentés de façon principalement théorique. Les bases de données disponibles en effet ont encore peu de recul temporel, comme nous le constatons sur les trois cas d'applications. Le problème du manque de données à observer (à l'origine de la méthode d'observation multi-niveaux) devient dans le cadre d'une observation cinématique plus prégnant, car il ne s'agit plus seulement de disposer de données, il faut également en posséder plusieurs versions. De fait, les possibilités de comparaisons entre les valeurs prises par une même information sont restreintes. Aujourd'hui, le manque de mises à jour est un frein important au suivi de phénomènes géographique dans un SIG.

B.3.2.2.2.1. Intégration de fonction d'interpolation entre deux valeurs calculées

Le premier mode d'enrichissement repose sur la densification des informations par intégration de valeurs estimées. Introduit dans le paragraphe B.3.2.1.3 pour assouplir le pas d'observation, il consiste à exploiter une fonction d'interpolation pour calculer des valeurs de caractéristiques entre deux versions. Il généralise ainsi l'utilisation implicite de la fonction identité (cf. figure B.3.11).

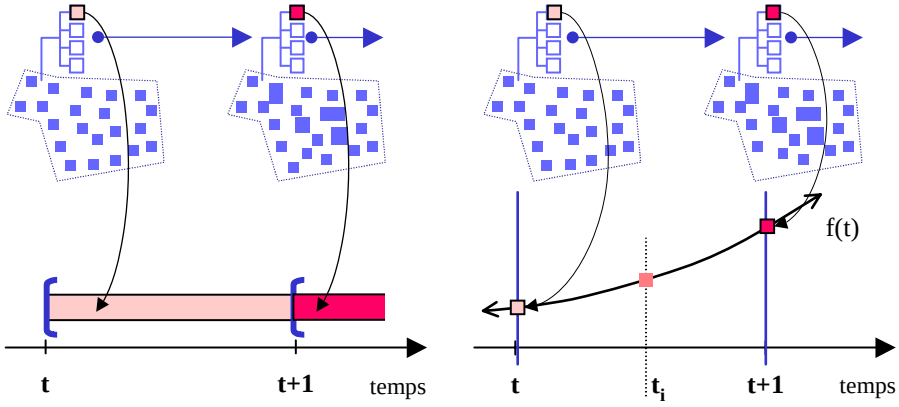


Figure B.3.11. *Utilisation d'une fonction d'interpolation entre deux valeurs calculées (fonction identité ou fonction plus élaborée)*

La difficulté de mise en œuvre dépend du type de phénomène et des caractéristiques à estimer (stables ou à fréquence de variations élevée, régulières / irrégulières). Elle dépend également des connaissances possédées par l'utilisateur sur le phénomène. En particulier, s'il dispose par ailleurs de modèles de simulation ou de projection, ces modèles trouveront là un champ d'application possible. Par exemple, à partir d'une description du terrain et d'un modèle de décrue, un utilisateur peut tenter d'estimer l'état de l'occupation des sols deux jours après un pic d'inondation. Ce genre d'applications peut être considéré comme un mode d'enrichissement de la méthode, car elles permettent de fournir des informations d'états intermédiaires (fig. B.3.12). Toutefois, les données à observer sont de moins en moins issues de l'observation directe du terrain, pour venir de plus en plus de la modélisation de ce terrain…

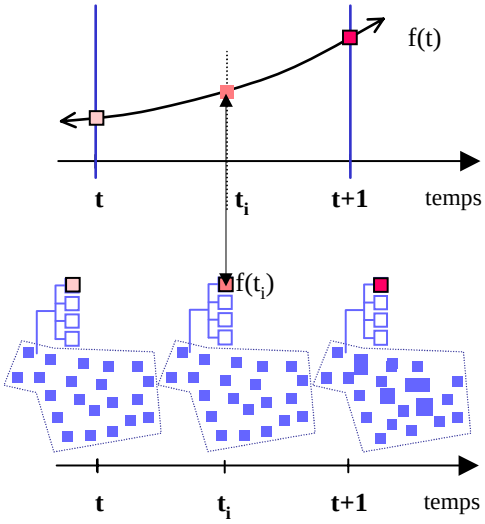


Figure B.3.12. *Intégration de valeurs de caractéristiques calculées par interpolation*

Cette utilisation des valeurs d'indicateurs comme jalon pour appliquer des modèles de simulation ouvre le champ à d'autres usages. En effet, ces valeurs utilisées comme points de repère peuvent servir à caler les modèles de simulation et à calculer ainsi des valeurs intermédiaires. Mais elles peuvent également servir dans le cadre d'une observation plus dynamique, pour tester de tels modèles et/ou valider des hypothèses d'évolutions. Ces applications apparaissent alors comme une extension de l'exploitation de la méthode, car elles l'intègrent à des analyses de phénomènes géographiques plus complexes. La méthode participe ainsi de moins en moins d'une utilisation des SIG comme *observatoire* pour s'orienter vers une utilisation de type aide à la *décision*.

B.3.2.2.2. Intégration de données supplémentaires utilisant un support d'agrégation différent

Respectant les conditions initiales d'application de la méthode, nous avons seulement envisagé des cas où les bases de données micro étaient compatibles avec une *même 'emprise'* d'agrégation (cf. fig. B.3.13) (ex. : les 'bâtiments' et les 'tronçons de route' de la BD TOPO®). Dans ce paragraphe, nous considèrerons le cas de bases de données micro compatibles avec des 'portions de territoires' différentes (c.a.d des représentations de l'*emprise* de l'*objet* observé différentes) (cf. figure B.3.14) : par exemple les données de la BD TOPO® de l'IGN et celles de BASE ILOT® de l'INSEE contiennent chacune une représentation distincte des 'îlots routiers'. Pour simplifier nous supposerons leurs dates de versions semblables.

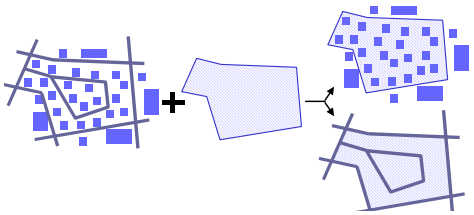


Figure B.3.13. *Exemple de données micro compatibles avec une même emprise*

Nous définissons un « système de données » comme *un ensemble de données micro compatibles avec une « même « représentation » de la portion de territoire »*. Nous proposons alors d'appliquer la méthode d'observation cinématique de façon autonome et indépendante à chaque « système de données » et de déduire des indicateurs pour décrire les agrégations d''objets micro' obtenues avec la représentation de l'emprise correspondant à chaque système.

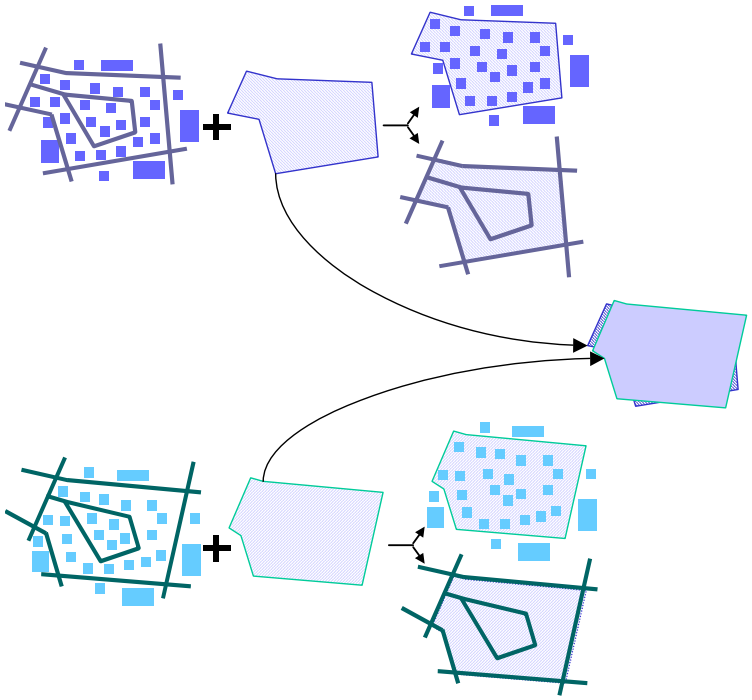


Figure B.3.14. *Exemple de « système de données » différents*

La même information, ici la maille du réseau routier, peut être représentée par des données différentes, car issues de processus de numérisation distincts (par ex. : les îlots de BASE ILOT ® de l'INSEE et ceux de la BD TOPO ® de l'IGN)

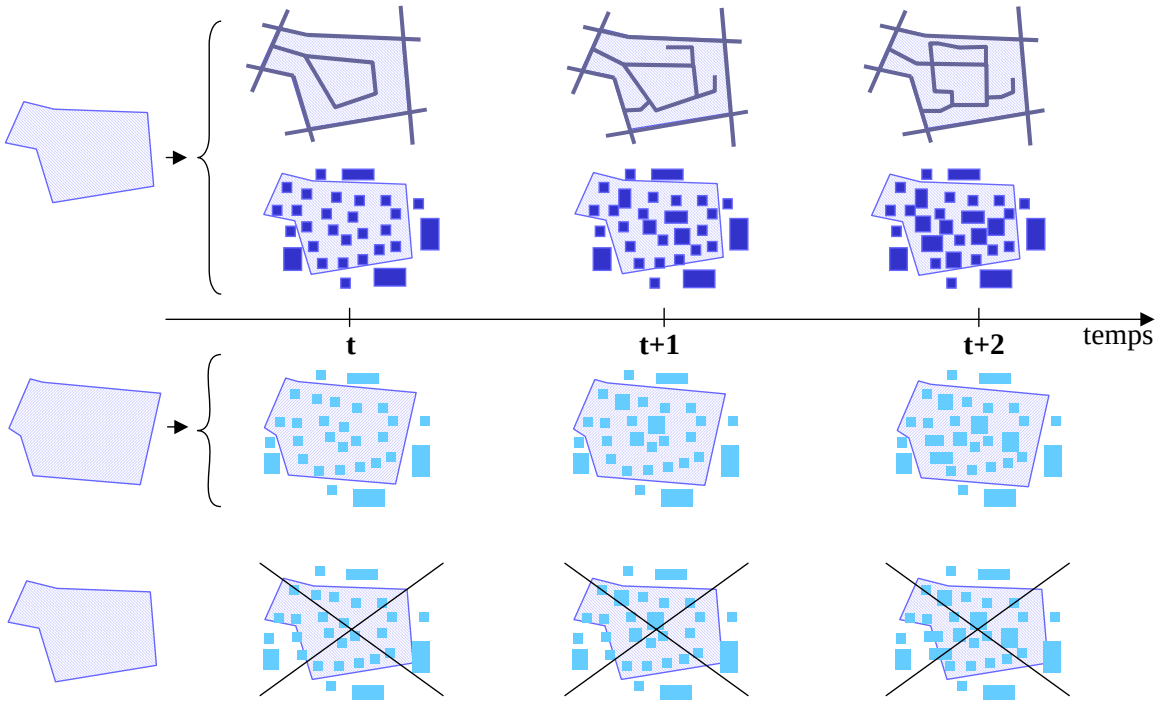


Figure B.3.15. *Utilisation d'emprises d'agrégation différentes*

Nous proposons de ne pas utiliser une seule ‘emprise’, mais d’exploiter des ‘emprises’ d’agrégation différentes pour assurer une plus grande compatibilité géométrique et une meilleure cohérence géographique (cf. figure B.3.15). Le principe de la méthode d’observation multi-niveaux proposée consiste a utiliser une même *emprise* de référence pour transférer des caractéristiques calculées sur des ‘agrégations d’objets micro’ à l’‘objet’ d’observation principal. Pour faciliter la présentation de la méthode et nous concentrer sur les aspects temporels de l’observation cinématique, nous avons traduit cela dans la pratique par l’unicité de l’‘emprise’. Toutefois nous avons vu avec la notion de *portion de territoire* (cf. B.2.7) qu’il est parfois préférable d’utiliser des ‘représentations’ différentes de l’*emprise* pour agréger les différents types d’‘objets micro’ (cf. figure B.3.16).

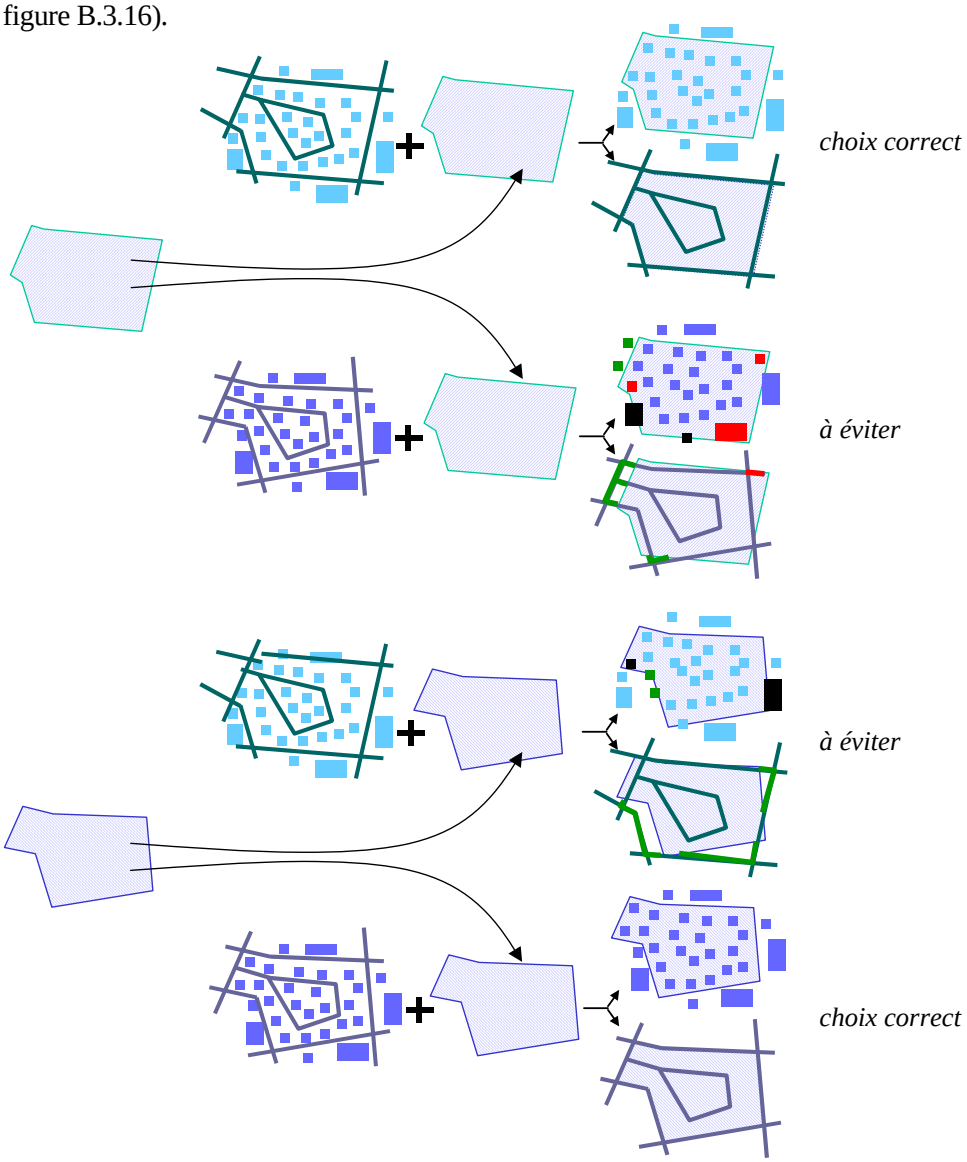


Figure B.3.16. *Cohérence de l’agrégation en fonction de l’emprise choisie*

Pour limiter les problèmes de cohérence entre des représentations différentes, il importe d’utiliser une représentation de l’emprise compatible avec le système de données auquel appartiennent les données micro à agréger.

Les indicateurs ainsi conçus décrivent la même *information*, objet méso géographique, mais sont portés par des composantes géométriques différentes. L’observation et l’étude conjointes de ces indicateurs dans un SIG pose alors un problème de structure de gestion, car les logiciels SIG ne permettent pas aujourd’hui d’associer à une même information géographique deux représentations différentes conjointes. Lorsque des

représentations différentes existent, elles sont gérées de façon indépendante, comme dans des calques (c'est le principe du *multi affichage*). La mise en œuvre dans un SIG d'une telle observation requiert donc un travail de modélisation et de conception préalable (cf. figure B.3.17). Par exemple, elle implique de répondre à des questions comme : Quelle doit être la composante géométrique attribuée à l'objet méso géographique observé ? Si on choisit une de celles associées à un « système de données » (par exemple celle ayant la meilleure qualité géométrique, celle supportant le plus d'informations,…) comment gérer les caractéristiques issues des autres systèmes ? Comment établir des liens ?… Indépendamment de la structure d'accueil, le regroupement automatique des caractéristiques décrivant une même portion de territoire soulève un autre problème : la nécessité d'apparier les 'portions de territoire'. Suivant les cas, cette opération sera plus ou moins complexe et coûteuse : de la simple jointure de fichiers sur la base d'un identifiant géographique (ex. : n° de parcelle, n° INSEE) à la mise en œuvre de techniques d'appariement utilisant des développements ou des « travaux manuels »[16].

16. On notera ici l'avantage de l'utilisateur qui peut adopter une base de données unique comme « support de référence ». En effet, l'intégration d'une grande variété de données sur un même support, par géocodage direct (à l'adresse ou par numérisation, sous forme d'information par nature) ou indirect (sous forme d'information localisée intégrée à la composante sémantique d'objets existants) lui garantit la cohérence des localisations relatives des objets pour une gamme d'utilisation (d'observation) fixée.
Ainsi, relativement à la méthode d'observation multi-niveaux, il évite le problème de modélisation complexe, car l'unicité du support lui permet de travailler avec un seul « système de données ».

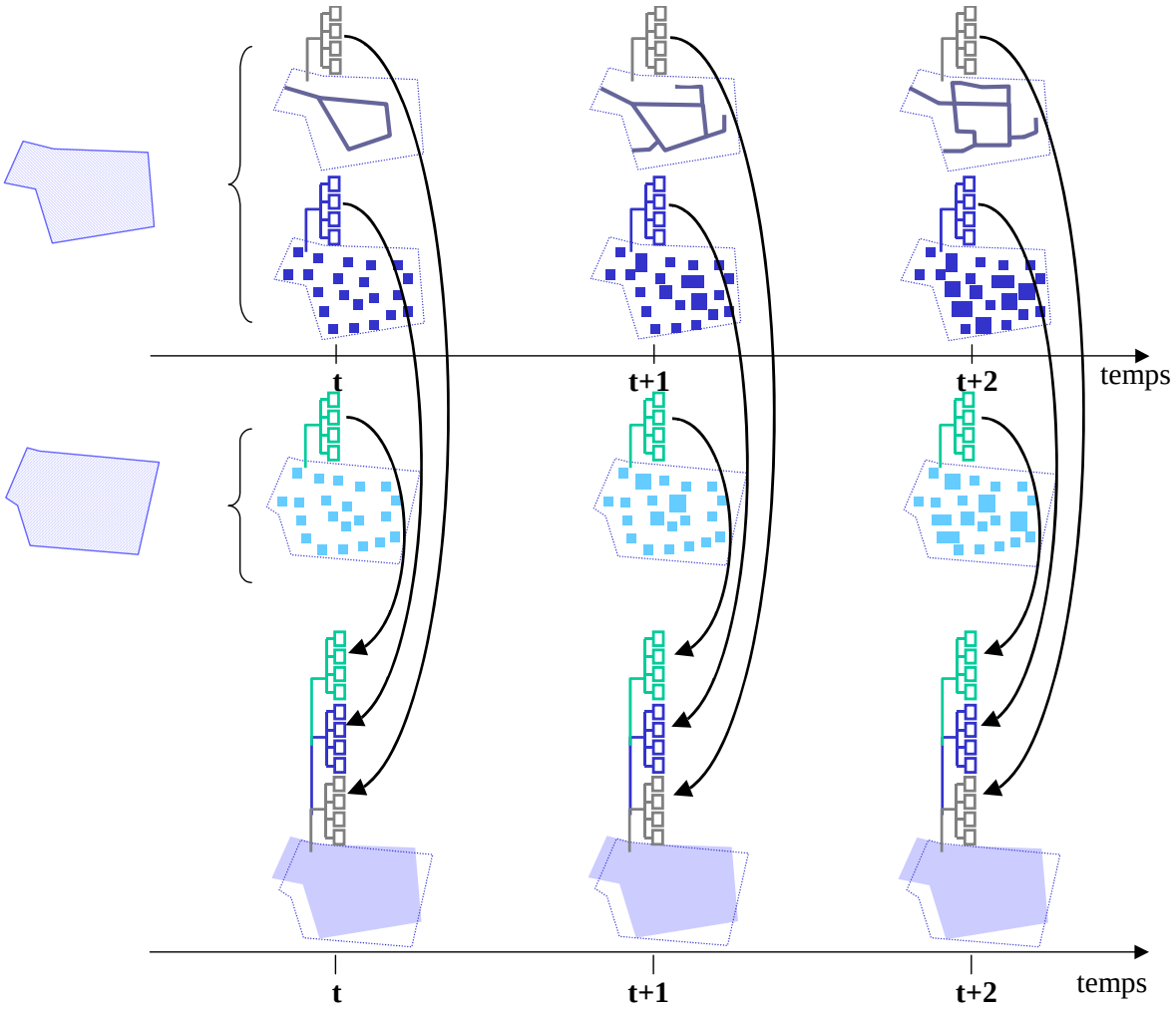


Figure B.3.17. *Intégration d'informations produites avec des « systèmes de données » différents.*

On peut exploiter dans une même observation des informations produites avec la méthode multi-niveaux à partir de données appartenant à des « systèmes de données » différents. En revanche, pour les intégrer dans un SIG, un travail supplémentaire de modélisation doit être fait pour définir quel sera le support de cet ensemble de caractéristiques.

B.3.2.2.2.3. Intégration de moyens de gestion d'informations temporelles plus sophistiqués

L'introduction d'un modèle de gestion des mises à jour plus sophistiqué permet de suivre plus commodément un objet dans le temps. Il facilite

l'extraction des couples d'états successifs et simplifie ainsi l'observation cinématique de l'objet étudié (cf. figure B.3.18a). Les outils d'appariement et de calculs différentiels facilitent quant à eux l'extraction automatique des changements entre deux états (fig. B.3.18b). Ils permettent :

- de construire une *représentation cinématique de l'objet à partir de mutations*, en qualifiant les mutations opérées entre les états successifs de l'objet[17]. Toutefois, la description des mutations ainsi obtenue reste souvent succincte : Elle se limite généralement aux changements élémentaires recensés dans la typologie des changements utilisée (fig. B.3.18c.);

- d'aider à obtenir des couples d'états successifs de caractéristiques pour une observation cinématique plus détaillée (ex. : l'identification uniquement d'un changement de type « modification sémantique » assure l'utilisateur qu'il y eu maintien de la composante géométrique) (fig. B.3.18d).

Jusqu'à présent, nous nous sommes appuyés sur le modèle par *snapshots* pour reproduire dans le temps des observations statiques et effectuer une *observation cinématique par état*. Dans ce paragraphe nous allons étudier l'apport de modèles et d'outils plus sophistiqués, en distinguant deux cas :

– Les moyens concernant des données méso,
– Les moyens concernant des données micro.

Auparavant, le choix d'une typologie des changements élémentaires doit être posé. Il explicite en effet les informations de changement auxquelles on accède directement grâce à ces moyens supplémentaires.

17. On retrouve ici les deux représentations duales pouvant être utilisées pour une observation cinématique: à partir des états, à partir des mutations (cf. A.2. P. 44).

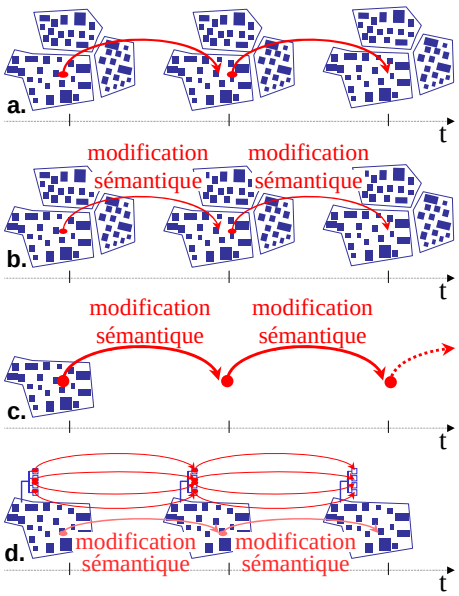


Figure B.3. 18. *Exploitation de moyens de gestion temporelle élaborés*

Choix d'une typologie des changements opérés sur les informations

Plusieurs typologies des changements peuvent être envisagées. Nous avons montré le risque avec une typologie très détaillée d'une interprétation erronée et le peu d'informations temporelles fournies par une typologie basée uniquement sur les données (cf. p. 170). Dans ce qui suit, nous avons fait le choix de travailler avec une typologie des changements restreinte[18]. Nous partons des composantes géométrique et sémantique d'une information géographique (cf. figure B.3.19a) et, dans une approche technique, nous définissons 5 catégories de changements (en fonction des changements élémentaires pouvant intervenir sur ces composantes) (cf. figure B.3.19b) :

- La *création* : une nouvelle information est introduite ;

- La *destruction* : une information existante est supprimée ;

- L'*identité* : une information existante reste la même en tout point sur l'ensemble de ses caractéristiques (comme dans la figure B.3.19b, où rien ne semble avoir changé sur les quais depuis le début du siècle) ;

- La *modification sémantique* : une information existante subit des modifications sur sa composante sémantique (ex. : changement de la valeur de l'attribut 'fonction du bâtiment' qui de 'gare' devient 'musée')

- La *modification géométrique* : une information existante subit des changements sur sa composante géométrique (ex. : changement de forme d'un bâtiment). La figure B.3.6 présente un cas limite de modification géométrique où l'ensemble du bâtiment est détruit sauf la façade.

18. …une typologie plus détaillée que la typologie effective sur les données {destruction / création} et plus factuelle que celle qui regroupe des changements élémentaires pour présenter des évolutions plus complexes (scission, fusion, agrégation).

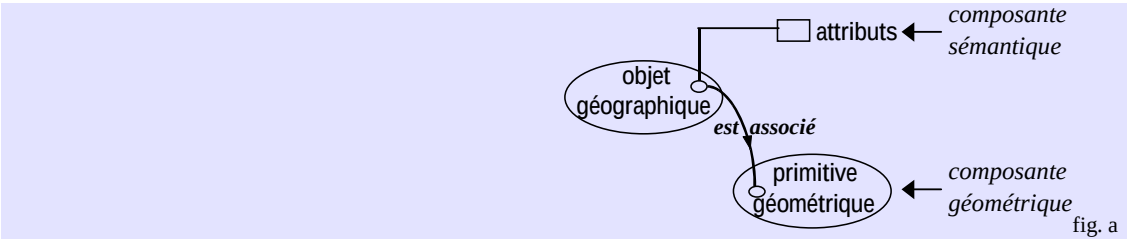


fig. a

Exemples de destruction/création (la Tour Montparnasse, la FNAC)

Exemple d'identité (les quais de Seine)

Exemple de modification sémantique (la gare d'Orsay)

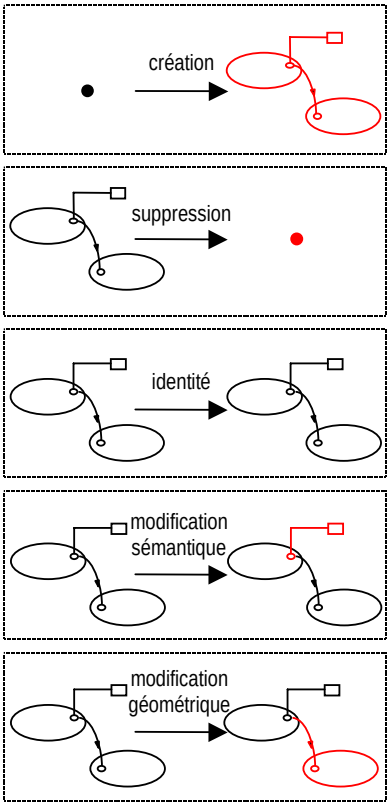


fig. b

Figure B.3.19. *Cinq types de changements sur les informations géographiques*

Informations temporelles sur les données méso

Supposons qu'indépendamment des données obtenues par la méthode que nous avons développée, l'utilisateur possède une base décrivant l'objet observé au niveau méso, pour laquelle il bénéficie de moyens de gestion temporelle sophistiqués. Il est alors capable de suivre facilement l'évolution de l''objet méso' dans cette base et de suivre ses caractéristiques dans le temps. Il peut donc les observer conjointement avec les caractéristiques calculées par la méthode proposée. Ainsi, l'utilisateur se trouve en possession de deux descriptions temporelles du même objet portant sur des niveaux de descriptions différents mais complémentaires : Une description méso intrinsèque à l'objet provenant de la base, une description résultant de l'agrégation des objets micro. Alors dans une analyse statistique dite *multiniveau*[19], l'utilisateur peut étudier les liens existants entre ces différents indicateurs et regarder par exemple, la contribution des caractéristiques issues des données micros aux attributs méso (et aux évolutions de ces attributs) (cf. figure B.3.20).

19. On retrouve ici le lien entre le *multi-niveaux* dans un SIG et le *multi-niveaux* en statistiques. Il ne s'agit pas d'un « simple glissement sémantique » (cf. p. 123), mais d'une véritable complémentarité pour l'analyse de phénomènes géographiques dans un SIG.

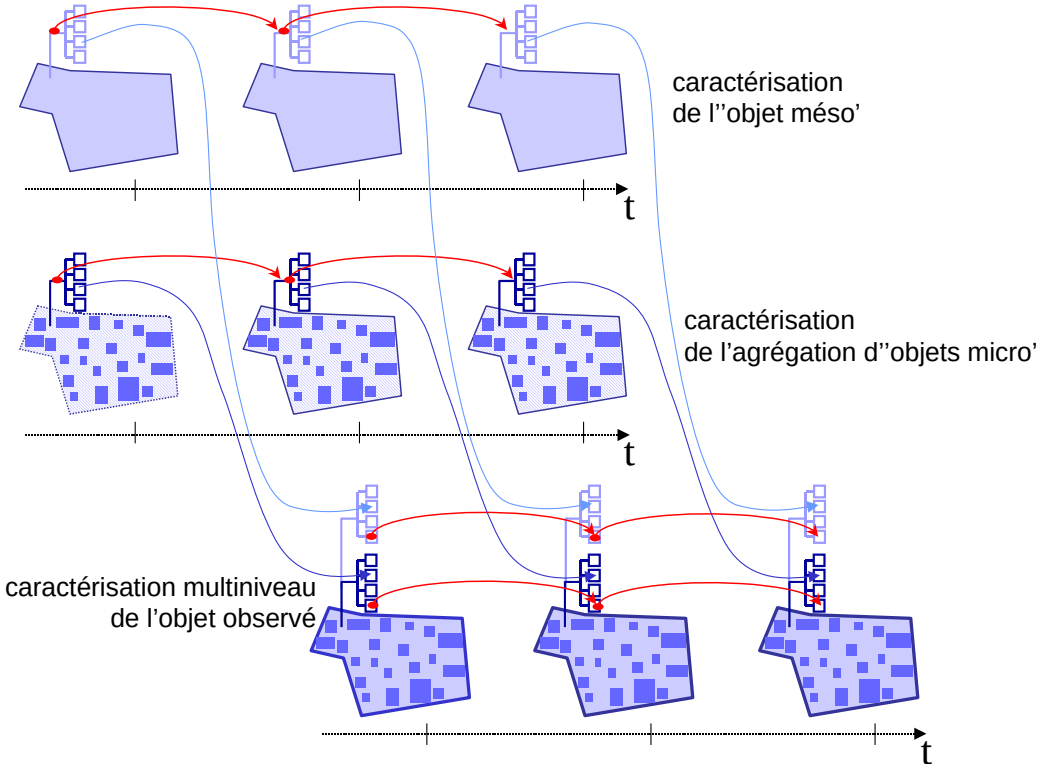


Figure B.3.20. *Observation multi-niveaux temporelle*

Il existe deux modes d'observation cinématique : par états, par mutations (cf. A.2. p. 44). Dans les SIG, les informations observées sont essentiellement des informations d'états, c'est à dire décrivant un objet (ou une caractéristique) à un moment donné. Nous ne possédons quasiment pas d'informations sur les mutations, mis à part les informations de mutations fournies au travers des couples d'états (sous la forme indirecte du différentiel entre ces deux états) et les informations de changements résultant de l'utilisation des outils différentiels… C'est un fait, **les modèles de gestion temporelle ne modélisent pas les 'objets temporels' « évènements »**[20] ; Les outils d'appariement et de calculs de différentiels ne permettent de saisir que des informations sur leurs effets et de façon globale, sans distinguer les effets provenant d'un événement, des effets dus à un autre. Dans les SIG, les moyens techniques les plus performants permettent au mieux de décrire l'action cumulée des évènements, à travers des changements opérés sur les données et selon une typologie restreinte des changements. Bref, ils ne permettent de saisir :

- aucun détail sur le *comment* ces évènements se produisent et la façon dont ils opèrent sur les objets géographiques,
- aucune information sur le *pourquoi* ils se produisent,
- aucune information même sur leur identité (*le quoi)*. En effet, on ne les saisit pas, on observe leurs effets. On ne les distingue même pas les uns des autres, les effets observés sont cumulés.

Seules les métadonnées peuvent actuellement accueillir quelques informations à leur propos (cf. A.3, p. 86).

Sans développement sur les schémas conceptuels utilisés par les SIG, l'observation cinématique ne peut porter que sur la trajectoire spatio-temporelle des 'objets' sous forme de *succession d'états* et non *pas de succession de mutations*. On ne peut pas envisager suivre les évolutions de l'objet observé sous l'effet des évènements. On ne peut accéder à la cause des changements, à leur origine… avec pour conséquence de limiter les *observations dynamiques*.

20. Comme nous le notons dans le paragraphe A.3.2 p.70, l'objet temporel « événement » n'a pas de modélisation explicite, on ne peut donc pas introduire d'informations pour le décrire. On ne peut connaître que l''état' sur lequel il opère et l''état' qu'il produit.

Tout en restant dans le cadre strict de cette méthode, l'intégration au niveau micro de moyens de gestion temporelle plus sophistiqués peut cependant apporter des éléments descriptifs sur les mutations d'objets. Dans ce qui suit nous allons voir qu'elle permet en effet d'enrichir la description des changements de l'objet méso en donnant des détails sur leurs formes (sur le « comment » (de quelles façons) ils se produisent ?).

Informations temporelles sur les données micro

On peut suivre les évolutions d'objets au niveau méso grâce à des modèles temporels (de type historique ou versionnement) ou grâce à des outils d'appariement. On peut suivre les évolutions de caractéristiques d'objets méso grâce à des outils de calculs différentiels et à la méthode d'observation multi-niveaux. Toutefois, il existe des cas où l'objet méso apparaît comme stable et où les agrégations d'objets micro paraissent constantes, donnant alors l'impression à l'observateur que tout est identique… alors qu'une observation au niveau micro montre des situations très différentes (fig. B.3.21).

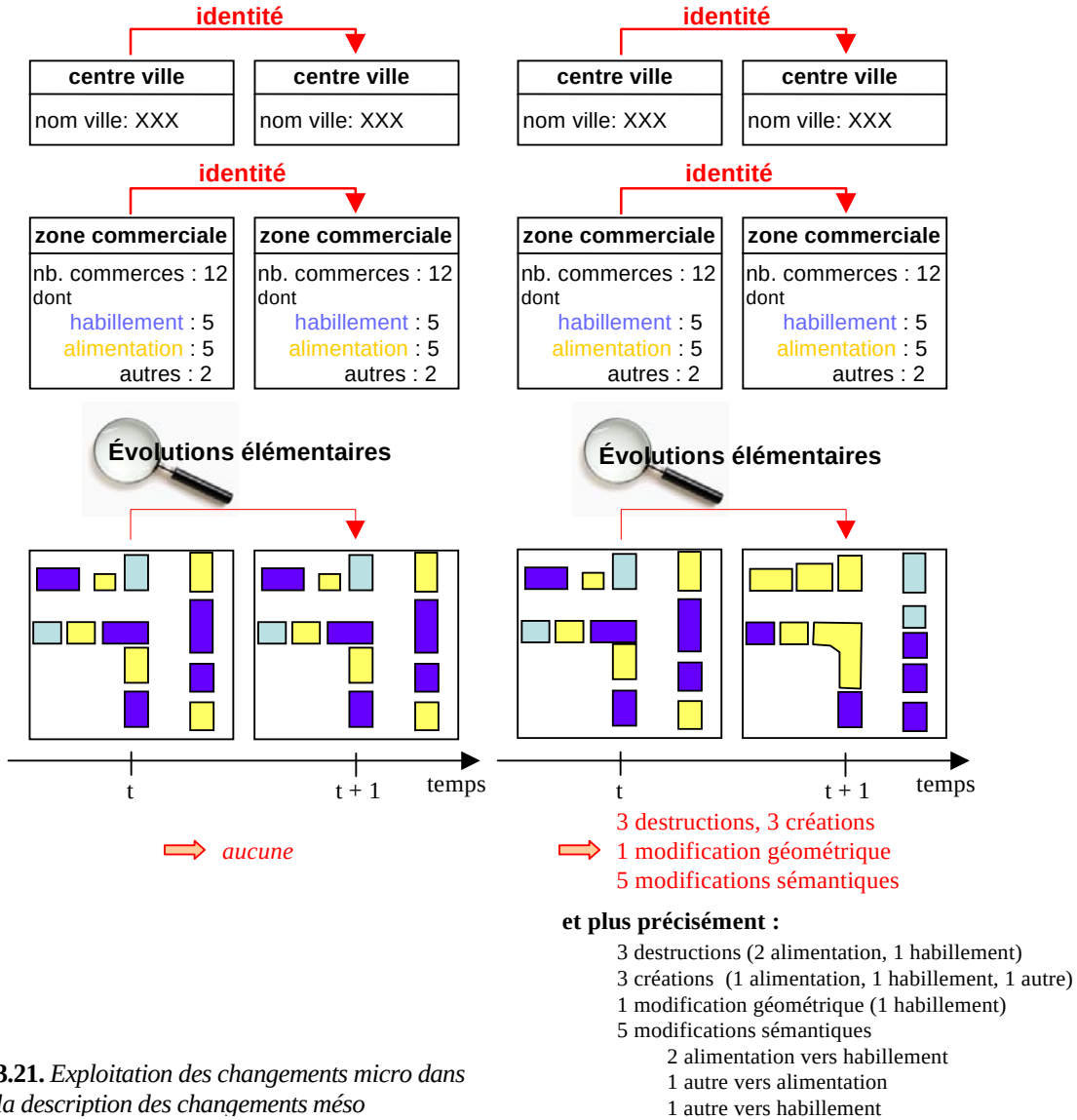


Figure B.3.21. *Exploitation des changements micro dans la description des changements méso*

L'exploitation des changements micro permet de distinguer des évolutions plus subtiles. Lorsque aucun changement au niveau méso n'est observable, elle sert à préciser s'il s'agit d'une « vraie » identité ou seulement d'une « identité apparente » et à détailler le cas échéant les évolutions micro. Ces observations supplémentaires, permises par l'utilisation conjointe de la méthode proposée et de moyens techniques de suivi temporel, ouvrent des perspectives pour l'observation de phénomènes émergents. Elles donnent également des éléments pour des études plus dynamiques. En effet, non seulement ces observations aident à déterminer s'il y eu évolution et à préciser son type, mais elles permettent également d'introduire d'autres informations descriptives sur ces évolutions, sur le « comment » elles se sont produites.

Considérons l'ensemble des changements micro ayant affecté l'objet méso observé sur une période fixée. Pour obtenir cet ensemble, il faut sélectionner les objets micros contenus dans l'emprise de l'objet et récupérer les changements micro intervenus sur ces objets pendant la période de temps considérée grâce aux moyens de gestion temporelle élaborés[21]. On peut alors caractériser cet ensemble :

- en calculant l'effectif de chaque changement micro élémentaire (identité, création, destruction, modifications sémantiques et / géométrique),
- en inventoriant les couples d'états d'objets micro ayant subis ces changements (par exemple : nombre d'objets micro ayant subis une modification sémantique les transformant de magasin d'alimentation en magasin d'habillement),
- en exploitant les dates des changements élémentaires micro pour les situer dans le temps et les uns par rapport aux autres. Cela permet par exemple de préciser la façon dont les changements micro ont contribué au changement au niveau méso (par exemple, la densification d'une zone pavillonnaire sur dix ans, s'est-elle effectuée de façon progressive tout au long des dix années ou bien a t'elle eu lieu d'un coup les premières années ou les dernières années ?)…

21. Par analogie entre l'espace et le temps, on peut considérer cet ensemble comme une représentation sous forme d'agrégation d'objets-changements micro, de l'objet-changement méso intervenu sur l''objet' méso observé.

Les informations de changements élémentaires micro permettent ainsi de créer des informations sur le *comment ?* se produisent les changements méso (par exemple, une 'zone pavillonnaire' peut rester identique par le nombre de 'bâtiments' qu'elle contient soit parce que ses bâtiments n'ont pas été modifiés, soit parce que le nombre de créations est égal au nombre de destructions). Dans ce mode d'enrichissement les descriptions d'états restent les mêmes (aucune caractéristique d'objet supplémentaire n'est produite) ; c'est la description des changements qui est enrichie (on peut discriminer les identités parfaites, on peut savoir combien d''objets micro' et quels 'objets micro' ont participé au changement de valeur d'un indicateur méso obtenu à partir de l'agrégation. Par exemple combien et quels 'bâtiments' ont contribué à l'augmentation de la surface moyenne du bâti sur une 'zone'). Cet enrichissement rend plus opérationnel le concept d'*objet suffisamment lui-même* en distinguant des évolutions plus subtiles, mais il ne permet pas de saisir ce qui n'est pas modélisé dans les SIG : « l'évènement »! En effet, relativement aux entités « évènements », les informations sur les changements fournies restent globales : On ne peut toujours pas distinguer les effets d'un évènements particulier ; Les changements micro dus à un même événement méso (c.a.d. se produisant au niveau méso) ne peuvent pas être regroupés automatiquement. (Par exemple, parmi tous les changements opérés sur un réseau routier, il ne sera pas possible d'identifier ceux correspondant à l'implémentation d'un nouveau rond point)[22].

22. sauf à prévoir des informations pour cela dans les métadonnées et les y avoir introduites au moment de la production des mises à jour.

En choisissant de maintenir l'observation de l'objet au niveau adéquat d'observation dit méso, et en « projetant » les informations micro sur ce niveau, nous permettons à l'utilisateur de disposer directement des informations qui l'intéresse au niveau qui lui convient. Ces informations sont compatibles avec les autres observations au niveau méso que l'utilisateur peut avoir mis en œuvre. Par ailleurs, ces informations sont simples à suivre dans le temps et permettent donc une observation temporelle dans un SIG. Enfin, si l'utilisateur dispose de moyens techniques sophistiqués pour saisir des évolutions au niveau micro, alors la méthode multi-niveaux lui permettra d'établir une description plus détaillée des évolutions méso.

C. Cas de mise en œuvre et d'application

C.1. Cas de phénomènes géographiques

Pour illustrer notre travail et montrer la diversité des phénomènes auxquels nos recherches peuvent s'appliquer, nous avons choisi trois exemples d'application. Il s'agit de *l'urbanisation pavillonnaire, de l'imperméabilisation des sols* (deux phénomènes particuliers liés à l'urbanisation) et du phénomène d'*ambiance urbaine*. Les cas sélectionnés relèvent principalement de l'urbanisme et de l'aménagement du territoire. Toutefois, ces recherches ne sont pas restrictives et nous aurions pu désigner d'autres cas pris dans d'autres domaines d'applications. En effet, ces travaux visent à proposer une *méthode générique* permettant de construire des *méthodes spécifiques* pour observer dans un SIG des phénomènes géographiques particuliers. La *méthode générique* est indépendante des champs thématiques considérés. Les cas d'applications qui ont été choisis servent avant tout à l'« instancier ». Ils permettent de tester l'applicabilité de la méthode et de valider l'intérêt des méthodes spécifiques qu'elle produit (fig. C.1).

Cette partie présente comment nous avons utilisé la méthode d'observation multi-niveaux pour élaborer des méthodes d'observation dédiées à chacun des phénomènes retenus, qui soient adaptées aux contextes d'applications particuliers à chaque cas (cf. figure C.2).

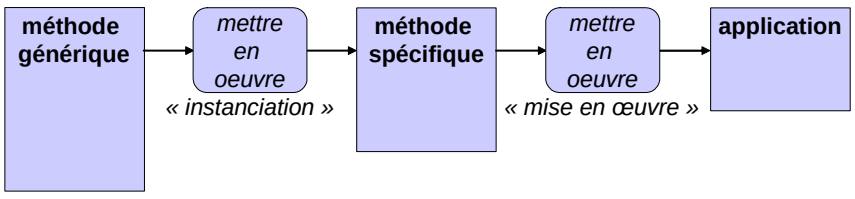


Figure C.1. *Méthode générique, méthode spécifique et mises en oeuvre*

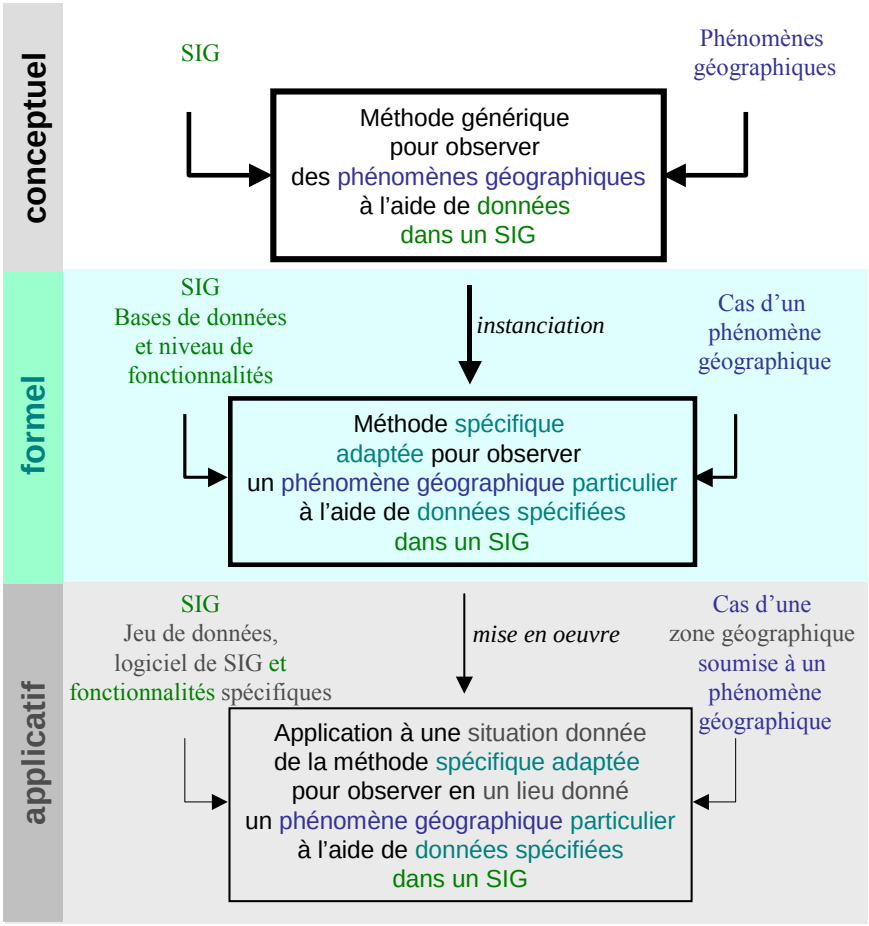

conceptuel
SIG
Phénomènes géographiques
Méthode générique pour observer des phénomènes géographiques à l'aide de données dans un SIG
formel
SIG
Bases de données et niveau de fonctionnalités
instanciation
Cas d'un phénomène géographique
Méthode spécifique adaptée pour observer un phénomène géographique particulier à l'aide de données spécifiées dans un SIG
applicatif
SIG
Jeu de données, logiciel de SIG et fonctionnalités spécifiques
mise en oeuvre
Cas d'une zone géographique soumise à un phénomène géographique
Application à une situation donnée de la méthode spécifique adaptée pour observer en un lieu donné un phénomène géographique particulier à l'aide de données spécifiées dans un SIG

Figure C.2. *Les différents niveaux méthodologiques : de la méthode générique aux applications en passant par les méthodes spécifiques*

Les trois cas d'étude forment un contexte représentatif des différents phénomènes géographiques que nous proposons d'observer avec un SIG.

- Le premier cas montre qu'il est possible d'améliorer les observations existantes sans impliquer de développements complexes. Il montre également comment introduire certains développements proposés par la méthode générique, pour faire évoluer la méthode d'observation mise en œuvre vers une méthode d'observation plus élaborée. L'application sur *l'urbanisation pavillonnaire* est la plus quantitative, mais aussi la plus complète. Elle met en œuvre l'ensemble de la méthode, sans privilégier l'une des étapes, en exploitant uniquement des fonctionnalités standards des SIG, tout en étudiant comment profiter de capacités d'analyse plus complexe et/ou de données complémentaires pour enrichir encore les observations. Elle met également en avant l'apport spécifique des analyses morphologiques et le caractère évolutif de la méthode d'observation.

- Le second cas met l'accent sur le travail analytique nécessaire à la construction des indicateurs de suivi, c'est à dire sur l'étape de *caractérisation*. Cette application qui traite du phénomène d'imperméabilisation, se concentre en effet sur la construction du *coefficient d'imperméabilisation*, indicateur donné par le ratio : surface imperméabilisée/surface totale. Conceptuellement, l'observation du phénomène d'imperméabilisation ne pose pas de difficulté, puisque les surfaces observées sont des *objets géographiques* facilement identifiables, accessibles par simple digitalisation de photographies aériennes. Dans ce cas, la difficulté réside dans le calcul de l'indicateur principal (le coefficient

d'imperméabilisation) et dans l'automatisation de ce calcul pour faciliter la mise à jour et son suivi. En effet, bien qu'accessible, la surface imperméabilisée n'existe pas sous forme de données dans les bases que nous avons exploitées. Il s'agissait d'exploiter des méthodes d'analyse pour la reconstituer.

- Enfin, le dernier cas oppose des conditions peu favorables à la saisie du phénomène étudié dans un SIG. Il s'intéresse en effet à un phénomène géographique immatériel, multi-dimensionnel et partiellement subjectif. Cette application met l'accent sur le travail conceptuel nécessaire à l'observation d'un phénomène dans un SIG. L'utilisation d'un tel outil implique en effet de disposer de données géographiques et donc d'avoir préalablement effectué un travail de modélisation en informations géographiques. Il s'agit donc d'identifier et de construire des informations pour représenter, caractériser et suivre le phénomène à étudier. Cette application détaille l'étape de *formalisation technique*. Elle illustre en particulier le travail de modélisation nécessaire pour disposer de l''objet méso-structurel' indispensable à l'application de la méthode d'observation multi-niveaux.

Ainsi, le premier cas nous permet de décliner entièrement la méthode générique et de montrer comment exploiter certains modes d'enrichissement proposés, tandis que les deux autres applications se concentrent sur les deux étapes clés de la mise en œuvre : la construction des indicateurs utiles au suivi du phénomène (la *caractérisation* B.2.5.3 p.146) et la modélisation de l'*objet d'observation principal* (la *construction* de l''objet méso-structurel' B.2.5.2 p.136).

C.2. Le phénomène d'urbanisation pavillonnaire

Le suivi de l'urbanisation pavillonnaire est un cas d'application important pour la méthode d'observation générique que nous proposons. En effet,

- il permet de mettre en œuvre l'ensemble de la méthode, de montrer ainsi son intérêt et comment elle fonctionne.

- il nous sert à illustrer de façon concrète et visuelle les principes fondamentaux de cette méthode, en particulier l'articulation entre l'étape d'observation statique centrée autour de la construction d'indicateurs de suivi et l'étape d'observation temporelle exploitant ces indicateurs (cf. Annexe 7)

- il permet de tester différentes conditions d'applications (en fonction des types de données géographiques disponibles et des compétences en SIG des utilisateurs). Il met ainsi en valeur le caractère générique de la méthode d'un point de vue technique.

- enfin, il sert à présenter certains enrichissements en matière d'intégration de données supplémentaires, d'exploitations d'outils et méthodes d'analyses, etc. Il montre en particulier, l'apport spécifique des analyses géométriques.

Les *zones pavillonnaires* qui s'imposent spontanément comme de « bons objets » pour suivre le phénomène d'urbanisation pavillonnaire, possèdent en effet de nombreuses qualités pour la méthode développée.

Conceptuellement, les *zones pavillonnaires* peuvent être appréhendées simplement comme des *ensembles de pavillon*. Visuellement, elles sont reconnaissables sur les photographies aériennes ou les sorties graphiques de bases de données à moyennes échelles, car elles apparaissent sous forme d'ensembles de constructions assez caractéristiques de par l'homogénéité de leur bâti et la régularité de leur structure morphologique souvent à plusieurs niveaux (cf. figures C.4 et C.5). L'exemple des *zones pavillonnaires* nous aide ainsi à présenter le travail analytique effectué dans la première partie de la méthode, qui vise à caractériser les 'objets' principaux d'observation pour construire ensuite des indicateurs utiles au suivi. Il permet également de montrer comment sophistiquer la méthode d'observation en intégrant des analyses plus élaborées pour enrichir l'étape de caractérisation. Dans cette application, nous proposons en effet d'introduire des analyses de la composante géométrique, pour mieux exploiter les différences morphologiques observées sur les images.

Par ailleurs, l'importance et la richesse de la dimension temporelle des *zones pavillonnaires* nous permettent d'illustrer le mise en œuvre de la seconde partie de la méthode dédiée à l'observation temporelle. En effet, bien qu'étant un phénomène récent, l'urbanisation pavillonnaire possède déjà une histoire (qui s'inscrit d'ailleurs fréquemment dans la structure morphologique de ses *zones*). Cette histoire peut être considérée à plusieurs échelles temporelles et spatiales (par exemple, l'évolution de l'urbanisation pavillonnaire peut s'envisager comme un mouvement général à l'échelle de plusieurs décennies ou au travers de changements plus localisés dans l'espace et dans le temps d'une *zone pavillonnaire* précise). Cette application offre donc à l'analyse plusieurs problématiques temporelles et aide ainsi à exprimer le potentiel de la méthode d'observation proposée.

Figure C.3. *Représentation cinématique de l'urbanisation pavillonnaire (bleu) sur une zone de Marne-la –Vallée (du haut vers le bas, en : 1982 ;1987 ;1990 ;1994 ;1999 [Extrait du MOS])*

Ces propriétés des *objets* géographiques *zones pavillonnaires* contribuent à faire de cette application un bon exemple pour présenter ces travaux. D'autres considérations, plus techniques cette fois, finissent de la transformer en exemple central pour cette thèse. L'étude du phénomène d'urbanisation pavillonnaire nous confronte en effet à des conditions d'application variées : des plus favorables où l'on dispose de données explicites sur les 'zones pavillonnaires' en tant que telles, au cas où l'utilisateur doit tenter de les reconstituer à partir de données sur d'autres 'objets' (par exemple des données sur le 'bâti') ; du cas précurseur où l'on dispose d'un modèle temporel sophistiqué permettant de retrouver les mutations (géométriques et/ou sémantiques) d'un objet, au cas où les données se présentent sous forme de *snapshot* laissant a priori la dimension temporelle à l'observation visuelle des utilisateurs.

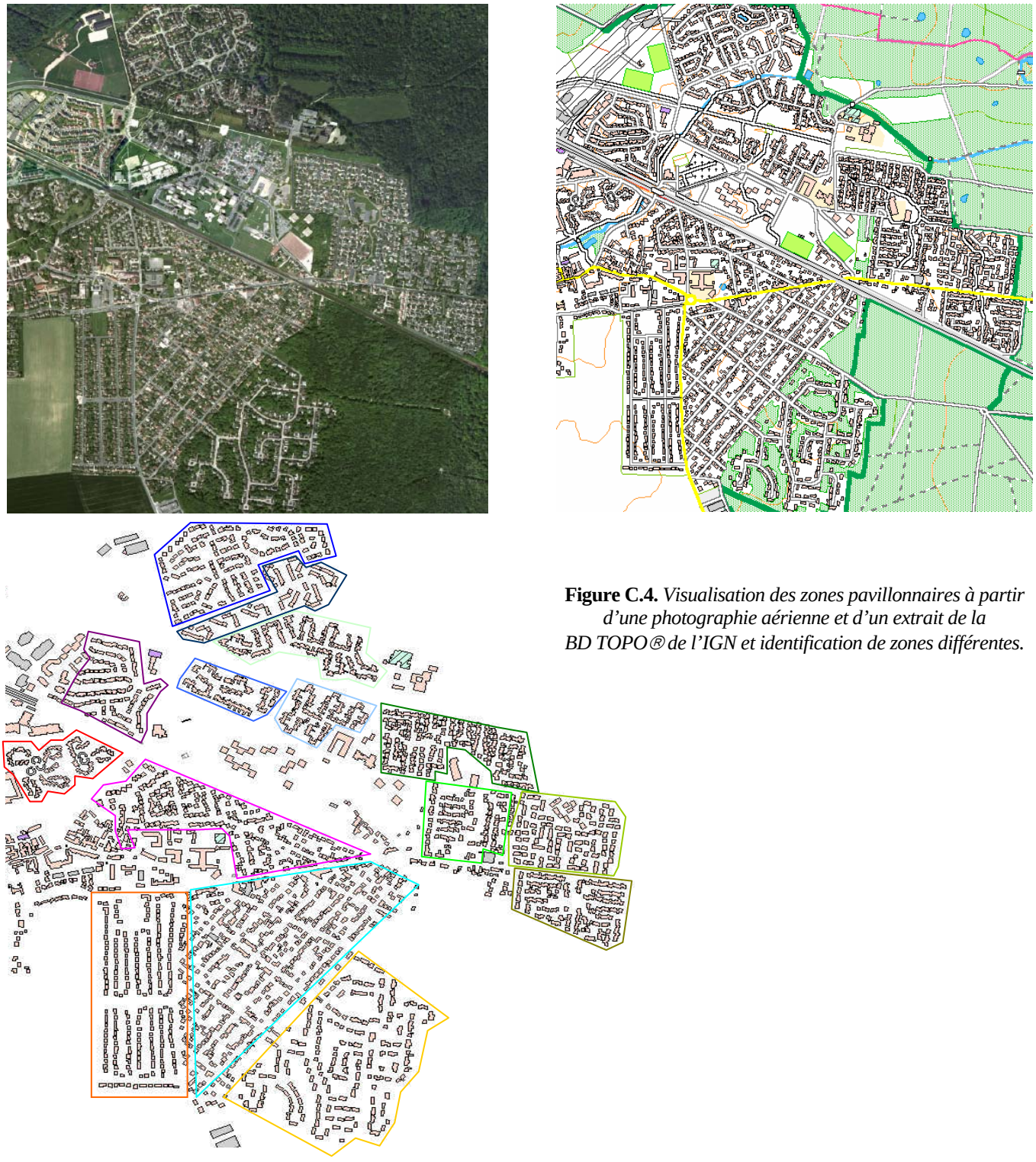
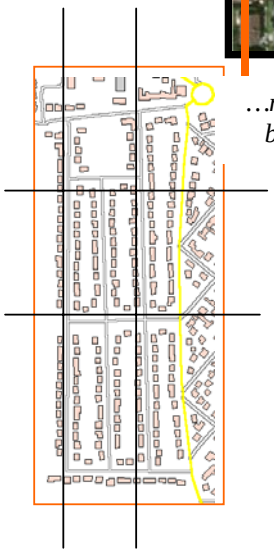


Figure C.4. *Visualisation des zones pavillonnaires à partir d'une photographie aérienne et d'un extrait de la BD TOPO® de l'IGN et identification de zones différentes.*

Trois exemples de zones pavillonnaires,
dont deux avec une morphologie des îlots très rectiligne…



…morphologie rectiligne que l'on retrouve au niveau du bâti de la zone orange mais pas de la zone turquoise.

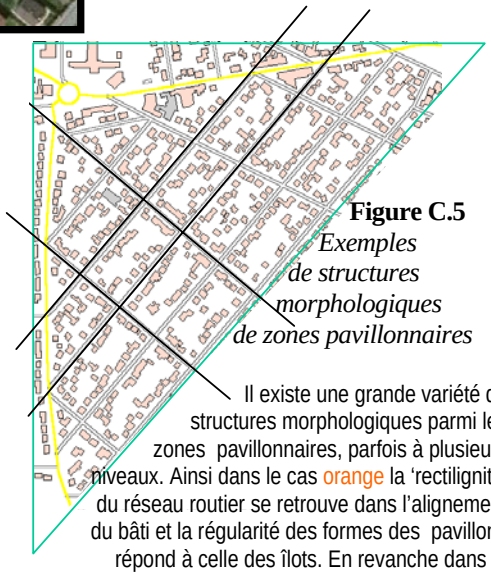


Figure C.5
Exemples de structures morphologiques de zones pavillonnaires

Il existe une grande variété de structures morphologiques parmi les zones pavillonnaires, parfois à plusieurs niveaux. Ainsi dans le cas orange la 'rectilignité' du réseau routier se retrouve dans l'alignement du bâti et la régularité des formes des pavillons répond à celle des îlots. En revanche dans la zone turquoise, ces caractéristiques n'atteignent pas le niveau des bâtiments.

C.2.1. Observations sur le phénomène d'urbanisation pavillonnaire

Apparue à la fin du 19ième siècle, l'urbanisation pavillonnaire est le mode d'habitat favori des Français[1] [Guerin, 1968]. Cependant dans le contexte actuel mettant en avant le thème général du développement durable et celui plus spécifique du renouvellement urbain, l'urbanisation pavillonnaire fait depuis quelques temps l'objet d'un discours critique de la part de certains aménageurs. Sans prendre part à ces débats, nous proposons de développer une méthode pour construire des indicateurs d'observation et de suivi de l'habitat pavillonnaire pour permettre aux décideurs de mieux les connaître. En effet, certains s'étonnent, au vu de l'importance de ce phénomène[2], du peu d'éléments dont ils disposent pour étudier l'évolution de ces zones et décider de leur avenir [Petit, 2000].

1. Enquête du groupe *De particulier à particulier,* publiée en janvier 2003 et réalisée sur un échantillon de 3800 personnes.

2. *« La forme urbaine que représente le pavillonnaire couvre des centaines d'hectares dans le nord de l'Essonne et souvent plus de la moitié, sinon les deux tiers du patrimoine bâti de certaines communes »* [Petit, 2000, p.3].

C.2.1.1. Observation à partir du MOS de L'IAURIF

Sur la région Ile de France, il existe une base de données, le MOS ou mode d'occupation des sols de l'IAURIF, qui intègre dans ses spécifications les objets 'zones pavillonnaires' en tant que tels. Elle permet de recenser et de localiser les 'zones pavillonnaires' à plusieurs dates entre 1982 et 1999 (cf. figure C.3). Par ailleurs, la mise en cohérence des partitions d'occupation des sols utilisées lors des différentes mises à jour et les choix opérés pour la structuration de ces données assurent les utilisateurs de pouvoir suivre ces zones dans le temps (cf. p.108). Ces données intégrées dans un SIG renseignent facilement sur la surface totale occupée par « le pavillonnaire », sur le nombre de 'parcelles pavillonnaires', sur leur superficie moyenne et sur leur localisation dans le territoire. Dans une perspective temporelle, cette base de données permet de retrouver les dates de changements d'occupations des sols, ainsi que le type d'occupation avant et après chaque changement.

Outre ces éléments d'observations statiques et cinématiques, le MOS nous a servi à montrer la pérennité des zones pavillonnaires. Dans une étude menée avec Hélène Mathian de l'UMR Géographie-cités, nous avons en effet observé que l'urbanisation pavillonnaire est un phénomène particulièrement pérenne à la fois dans ses effets (une *zone pavillonnaire* construite tend à perdurer) et en tant que processus (les *zones pavillonnaires* existantes s'étendent et de nouvelles sont régulièrement produites). Cette étude s'appuie sur l'analyse des couples (type d'occupation des sols de la parcelle n à la date t_i ; type d'occupation des sols de la parcelle n à la date t_{i+1}) pour n décrivant toutes les parcelles et i toutes les dates de mise à jour du MOS. Cette étude met en évidence la rareté des couples du type (état pavillonnaire à la date t_i ; état autre que pavillonnaire à la date t_{i+1}) d'une part, qui démontre la longévité des zones pavillonnaires[3]. D'autre part, elle montre que pour l'ensemble des dates t_i, il existe un nombre significatif de couples du type (état non pavillonnaire à t_i ; état pavillonnaire à t_{i+1}) attestant de la production de nouvelles zones (cf. Annexes 7). L'analyse de la répartition dans le temps de ces créations, sur un territoire plus large permettrait également d'identifier les phases de développement et de récession du phénomène.

3. Plusieurs raisons fournies par les aménageurs peuvent expliquer cette longévité {Monjal, 2003}. Ils l'expliquent par exemple par des raisons politiques (pour modifier un lotissement en vue de construire d'autres aménagements, il est souvent nécessaire d'exproprier les propriétaires, ce qui est rarement populaire), pratiques (les nouveaux programmes immobiliers nécessitent souvent d'importantes superficies de terrain, or les zones pavillonnaires sont souvent découpées en petit parcellaire, impliquant de négocier avec une multitude de petits propriétaires), sociales et économiques (pour acheter leur pavillon, les familles s'endettent pour de nombreuses années et elles y restent longtemps attachées), etc.

Toutefois, pour observer de façon comparative deux 'zones pavillonnaires' uniquement à partir du MOS, l'utilisateur ne peut exploiter que la composante géométrique de ces 'zones', la différenciation s'effectuant alors sur la base de la localisation et des caractéristiques géométriques de l'emprise au sol des zones. En effet, dans le MOS les 'zones pavillonnaires' sont représentées comme des *portions de territoire*

homogènes, caractérisées uniquement par leur type d'occupation des sols, sans autre information attributaire pour décrire ces *portions de territoire*. Ainsi dans le cas de la figure C.6, où l'on considère pour schématiser que les quatre 'zones' observées ont la même forme d'emprise, alors seule la localisation peut permettre de les distinguer. Or elles peuvent correspondre à des situations très différentes (tant du point de vue de la densité, que de la répartition du bâti sur l'ensemble de la surface ou encore du nombre de pavillons…).

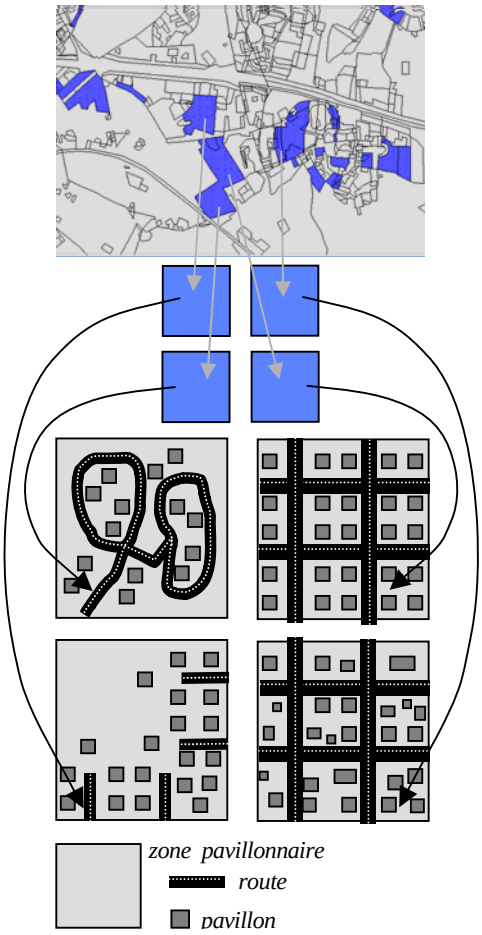


Figure C.6. *Vers une observation comparative des zones pavillonnaires*

Par ailleurs, l'utilisation du MOS dans le cadre d'une observation temporelle, donne principalement à voir des créations et des destructions de zones. Elle ne permet pas de saisir des évolutions différenciées, ni d'identifier les prémices aboutissant à l'émergence progressive de nouvelles zones. Dans le schéma de la figure C.7 on voit en effet qu'un même exemple d'évolution dans le MOS (a), où des zones sont recensées comme des 'zones d'espaces verts' dans plusieurs versions de la base et deviennent à une certaine date 'zone pavillonnaire', peut correspondre sur le terrain à des évolutions variées, à des processus de mutations différents : implantation rapide d'une zone entière (b), montée en puissance, densification lente (c) ou encore apparition insidieuse comme dans le phénomène de mitage (d).

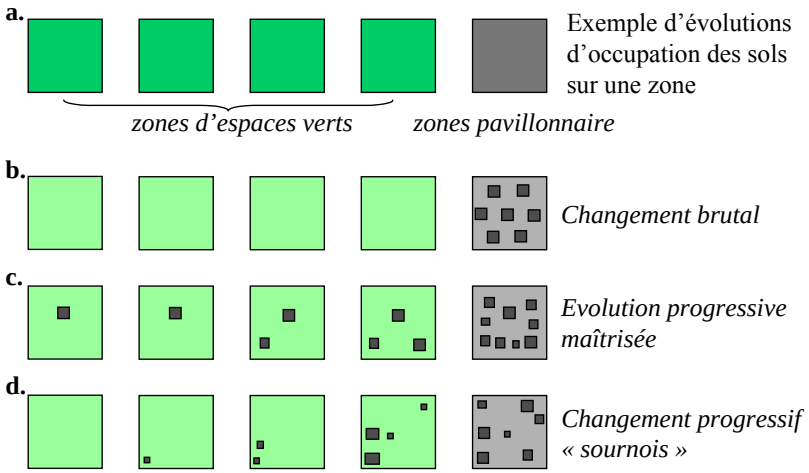


Figure C.7. *Une même évolution générale peut correspondre à des situations détaillées différentes*

Actuellement, l'implantation des lotissements correspond à un changement brutal (b) ; Un promoteur unique se charge de construire d'un « seul coup » tout un ensemble de maisons pour bénéficier d'économies d'échelle.

La montée en puissance des lotissements au début du 20ième siècle a suivi un mouvement d'évolution progressive (c) (que l'on retrouve aujourd'hui dans la densification de certaines zones) ; chaque propriétaire se charge de faire bâtir sa propre maison (les « maisons de maçons »).

Le mitage des espaces verts (des espaces boisés) correspond à un changement progressif « sournois » (d). Petit à petit, l'espace vert se transforme en 'zone d'habitat' par implantation de nouvelles constructions et « petites » transformations successives, souvent illégales. Ce phénomène est particulièrement préoccupant dans le sud de la France et tout le long du littoral {Rigail 2004}, {Garcez, 2005}.

Ces deux exemples d'observations montrent comment des informations géographiques sur les *bâtiments* peuvent servir à caractériser les *zones pavillonnaires* de façon complémentaire (par exemple à partir de leur densité, de leur homogénéité, de leur nombre de bâtiments, de la surface

moyenne du bâti, etc.). Ils montrent également qu'une caractérisation plus fine permet également de mieux différencier les zones et leurs évolutions.

C.2.1.2. Vers une caractérisation fine des zones pavillonnaires

Caractérisations thématiques

Pour effectuer une observation plus complète sur les zones pavillonnaires, il est utile de disposer d'informations plus détaillées. Les *zones pavillonnaires* ne manquent pas de caractéristiques thématiques. Les démographes savent par exemple, qu'il existe une relation privilégiée entre les familles avec jeunes enfants et l'accession à la propriété d'un pavillon avec jardin[4]. Les urbanistes s'intéressent également aux *zones pavillonnaires* en tant que mode d'habitat préféré. Ils observent les formes très particulières qu'il prend parfois (comme par exemple avec les *gated communities*[5]). La zone pavillonnaire est en effet occasionnellement un objet d'expérimentation (cf. la première photo de la figure C.8). Sans envisager des extrêmes, les sociologues ont observé l'influence de la structure morphologique dite en « raquettes » sur les phénomènes de ghettoïsation {AUE, 2004}. Dans un autre registre [Millot, 2003] a montré l'existence de corrélations entre le nombre d'accidents impliquant des enfants habitant une zone pavillonnaire et la structure de celle-ci… Pour traiter ce type de problèmes il faut avoir accès à des bases de données thématiques (ex. : données démographiques type INSEE, données sur les accidents type CONCERTO, etc.), mais également à des indicateurs strictement morphologiques du domaine des SIG. En effet, la structure morphologique des zones pavillonnaires apparaît ici comme un élément de caractérisation particulièrement intéressant à saisir, car elle participe à de nombreuses analyses thématiques.

4. Certaines collectivités ont d'ailleurs subi les conséquences de cette relation très étroite. Pour répondre aux besoins en équipements (crèches, écoles) générés par l'afflux massif de « population pavillonnaire », certaines collectivités ont investi dans l'aménagement de nouvelles structures. Ces structures n'étaient pas encore amorties qu'elles ne correspondaient déjà plus aux besoins, qui portaient alors sur d'autres structures (collèges, lycées), destinées, elles aussi, à subir le même sort { Morand, 2003}. Cela explique peut-être en partie la mixité de certains projets d'aménagement récents comprenant de l'habitat individuel et du collectif et permettant d'accueillir une population plus hétérogène.

5. Les *gated communities* désignent des quartiers dont l'accès est contrôlé, interdit au non résidents et dans lesquels l'espace public (rue, trottoirs, parc,...) est privatisé.

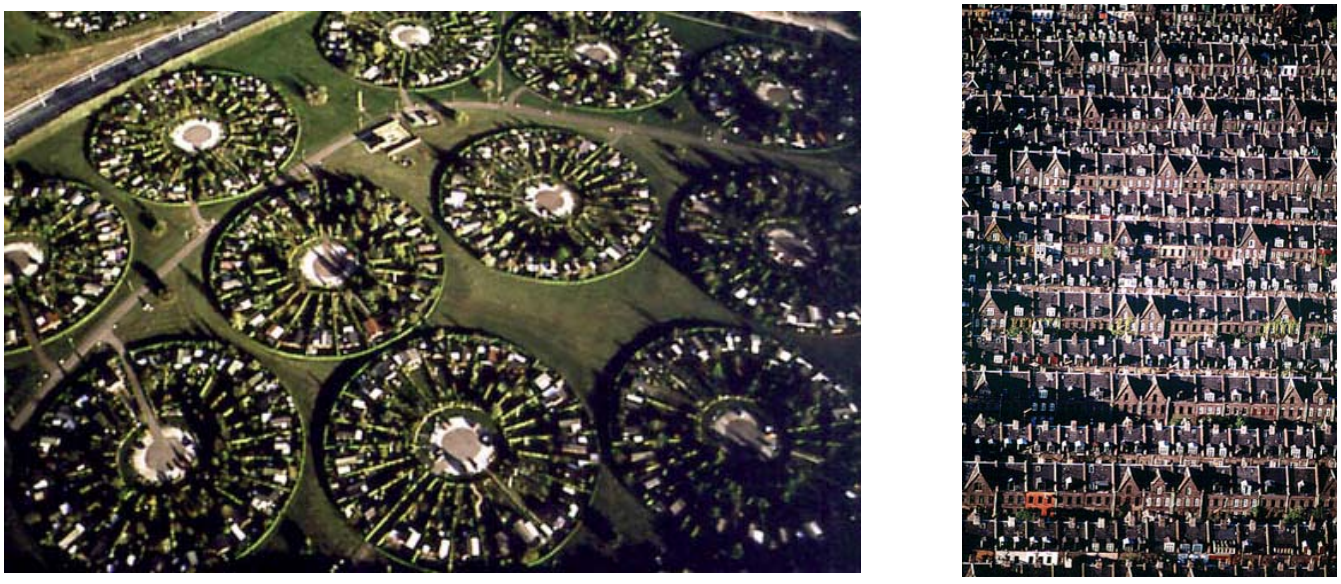


Figure C.8. *Deux exemples de zones pavillonnaires aux morphologies très marquées (Photo de Y. Artus-Bertrand)*

Caractérisation morphologique

La caractérisation morphologique des zones pavillonnaires pourrait contribuer à des études thématiques. Elle pourrait également apporter des éléments utiles pour l'observation du *phénomène d'urbanisation pavillonnaire* lui-même : elle servirait à construire des indicateurs spécifiques supplémentaires pour une observation plus discriminante. Ces indicateurs seraient d'autant plus pertinents que, comme l'illustre la figure C.8, la structure morphologique est un élément de différenciation

forte entre les zones. Cependant, pour saisir ces caractéristiques, les fonctionnalités standards des SIG ne suffisent pas ; Il est nécessaire d'intégrer des analyses propres à la composante géométrique. Après une mise en oeuvre élémentaire de la méthode, nous avons choisi d'intégrer de tels développements pour une application plus sophistiquée (cf. Annexe 7). Cette application a valeur illustrative : elle montre comment enrichir la méthode d'observation en introduisant des indicateurs morphologiques. De façon plus générale, ce type de développement montre comment améliorer les capacités d'analyses en diversifiant les indicateurs de suivi.

Par exemple, il est intéressant de noter que la structure morphologique des zones dépend généralement du style architectural et des principes urbanistiques ayant cours au moment de leur construction, les mouvements architecturaux subissant eux aussi des phénomènes de « modes ». Ainsi, sous réserve d'avoir établi un lien entre ces styles architecturaux et les indicateurs morphologiques, on pourrait associer aux zones un « style » en fonction de leurs caractéristiques et donc indirectement une époque de construction probable. Dans une observation statique du territoire, une telle caractérisation morphologique pourrait aider à faire une « lecture temporelle » des espaces pavillonnaires.

Dans une perspective plus dynamique, on peut s'intéresser aux différents types d'évolutions et chercher en particulier à distinguer :

- les zones « pérennes souples », c'est à dire celles qui tout en se maintenant, continuent à évoluer. En particulier celles qui continuent à se développer en se densifiant, éventuellement en subissant des micros changements, des extensions (ex. : garages, vérandas, pièces supplémentaires),
- les zones « pérennes rigides », qui se maintiennent strictement, à l'identique, sans accueillir de nouveaux bâtiments ou toute autre modification.

On peut regarder s'il existe une relation entre des évolutions différenciées et des morphologies particulières et tenter de répondre à des questions du type : *Certaines formes pavillonnaires sont-elles plus favorables à des évolutions ultérieures que d'autres ?* Par exemple, *certaines structures pavillonnaires sont-elles plus propices à la densification ?* Ces questions font écho aux préoccupations des aménageurs, lorsqu'ils s'interrogent sur le devenir des zones pavillonnaires ou quand ils abordent le thème de la « densification urbaine »… Toutefois, pour prendre des décisions, les aménageurs devront préciser en quoi la structure morphologie est une *causalité directe* et/ou *indirecte* d'un type d'évolution. Autrement dit, ils devront déterminer si certaines formes favorisent l'évolution ou si elles sont seulement corrélées aux facteurs qui l'expliquent. En effet, nous avons vu que la morphologie est fonction de l'époque de construction. Or l'époque de construction participe également en tant que paramètre à la probabilité de mutation d'un pavillon[6]. Les aménageurs devront donc identifier parmi les indicateurs morphologiques ceux qui sont des paramètres directement explicatifs de l'évolution et ceux qui n'interviennent que comme référent à une époque de construction (l'époque de construction possédant elle-même des caractéristiques pouvant influer sur les changements)[7].

C.2.2. Mise en œuvre de la méthode pour l'observation des zones pavillonnaires

Dans cette application nous avons décliné la méthode générique développée dans cette thèse, pour élaborer une méthode d'observation

6. Cette contribution est également indirecte, la date de construction étant révélatrice d'autres paramètres plus directement influents. Par exemple : l'age de l'accédant à la propriété et la durée de la propriété ou encore les conditions économiques à l'époque de la construction comme un prix du foncier bas… En effet, dans ces exemples, la date de construction symptomatise alors un bâti ancien qu'il s'agit de rénover, une étape significative dans la vie du propriétaire (départ du foyer des enfants ou entrée dans le grand age des parents) expliquant le *turn-over*, l'attractivité d'un marché foncier permettant une importante plus-value justifiant la cession d'une partie de terrain...

7. Par exemple, dans la note 4. nous évoquons le cas des zones construites récemment, « d'un coup », qui se caractérisent par une très grande homogénéité (elles offrent souvent des déclinaisons d'un unique produit) et qui attire en conséquence des profils familiaux très proches. L' évolution de ces zones se fera probablement par mutations massives. Ces vagues correspondront à des périodes où toutes les familles se trouveront en même temps, à traverser des étapes particulières de leur vie, favorables aux changements. (Pour identifier ces étapes, on pourra s'intéresser aux travaux en sociologie portant sur les parcours individuels). Est-ce pour amortir les pics de changement que certains programmes récents offrent une plus grande variété de logements, mélangeant parfois même collectif et pavillonnaire ? En effet, sans élément de la part des aménageurs, nous avons cependant constaté dans cette application, que l'hétérogénéité morphologique semblait être un paramètre favorable à la densification, contrairement à l'homogénéité.

dédiée au suivi de l'urbanisation pavillonnaire. Adoptant deux approches de complexité croissante, nous proposons deux mises en œuvre différentes.

C.2.2.1. Exploitation élémentaire de la méthode générique

La première mise en œuvre correspond à une exploitation basique de la méthode. Elle utilise des données géographiques existantes sur les objets 'zones pavillonnaires' en tant que tels[8] et exploite des fonctionnalités standards des SIG pour enrichir la caractérisation de ces 'zones' à partir de données sur les 'bâtiments' (dénombrement, moyenne sur la valeur d'un attribut, ratio entre les valeurs de deux attributs, buffer). On détermine par exemple : le *nombre de bâtiments* dans la zone, la *surface moyenne* de ces bâtiments, la *densité* du bâti, la *surface d'espace non bâti*. Dans un premier temps, les nouvelles caractéristiques permettent une observation statique plus détaillée et peuvent servir à la rédaction de cartes thématiques sur les zones pavillonnaires. Dans un second temps, nous calculons ces indicateurs pour deux versions différentes de la base de données sur le 'bâti'. Par comparaison, nous sommes alors en mesure d'identifier des évolutions relatives aux 'zones pavillonnaires' et de repérer, par exemple, celles qui se sont densifiées dans le temps (cf. Annexe 7).

Dans cette application, nous envisageons également le cas où l'utilisateur ne dispose pas de données sur les 'zones pavillonnaires' en tant que telles. Or, sans données sur ces 'zones', nous ne disposons plus de support auquel rattacher les indicateurs calculés. L'utilisation d'un SIG pour le suivi d'évolutions nécessite en effet d'associer ces indicateurs à une information géographique support, c'est-à-dire techniquement de les intégrer à la composante sémantique d'une information géographique possédant elle-même une composante géométrique explicite. Le chapitre B.2 nous permet d'envisager différentes solutions pour pallier ce manque de données :

- utilisation des 'îlots routiers' comme autres *données supports* pour agréger les 'bâtiments',
- utilisation des relations de voisinage pour construire des clusters d'objets géographiques auxquels associer les indicateurs.

C.2.2.2. Exploitation développée de la méthode générique

La seconde mise en œuvre introduit des éléments de sophistication au sein des deux étapes principales de la méthode qui consistent respectivement à :

– construire de nouveaux indicateurs d'observation statique,

– suivre l'évolution de ces indicateurs dans une observation cinématique.

Dans cette application, nous introduisons de nouveaux indicateurs : nous utilisons pour cela des analyses développées, qui intègrent en particulier des aspects morphologiques relevant du *bâti* et du *réseau routier*. Pour élaborer ces indicateurs, nous exploitons des données supplémentaires sur le 'réseau routier', et nous procédons à des analyses géométriques qui n'existent pas dans les logiciels de SIG. Les indicateurs calculés sur la forme du 'bâti' (concavité, élongation) et sur le type des 'tronçons de route' (rectiligne, impasse) nous servent ainsi à réaliser une typologie des zones pavillonnaires en fonction de leur style architectural (désordonné, ordonné, en grappe, groupé) (cf. figure C.9) (cf. Annexe 7).

8. Il serait en fait plus rigoureux de parler de « *zone d'habitat individuel* » plutôt que de « *zone pavillonnaire* » *(cf. annexe 7)*. Dans le MOS en particulier, il s'agit bien de données sur les 'zones « d'habitat individuel' » et non sur les 'zones pavillonnaires'. Cependant, dans cette présentation, nous avons préféré le terme de « pavillonnaire » pour rester proche de la perception de beaucoup de non-spécialistes de l'habitat, et de la perception collective (on entend plus souvent « *je rêve d'habiter un pavillon* » que « *je rêve d'habiter une maison d'habitat individuel* »). Nous espérons ainsi conserver à l'exemple sa valeur illustrative générale et l'ouvrir à des thématiques qui ne soient pas uniquement du registre spécifique de l'habitat ou de la construction, même si en toute rigueur :

Un pavillon est : « une habitation individuelle reculée par rapport à la rue, avec un jardin devant, un terrain derrière et une absence de contrainte mitoyenne dans la construction » [Raymond et al, 2002].

Des "maisons jumelées" forment un groupe aligné de deux maisons individuelles [Mangin et al, 1999]. Plus généralement, lorsque les maisons sont construites sans espace entre elles, formant un bâtiment unique, on parle de "maisons en rangée" ou de "maisons individuelles construites en bandes" [Garnier, 1984] ou d'"individuel groupé" [UNCMI]. Par la suite, nous désignerons par l'expression "maisons en rangée" tout ensemble de maisons individuelles accolées regroupant au moins deux maisons.

De façon générale, la "maison individuelle" désigne une habitation de type pavillon ou une maison intégrée à un groupe de maisons en rangée.

1) type « désordonné »

2) type « ordonné »

3) type « en grappe »

4) type « groupé »

*Dans les zones de type **« désordonné »**, une grande partie des pavillons ont été construits entre 1894 et 1954. Leur forme est variée et leur agencement dans l'îlot n'est pas organisé. Le réseau routier est rectiligne.*

*Les zones de type **« ordonné »** (1954-1970), sont généralement très uniformes, composées de pavillons de forme rectangulaire, desservis par un réseau routier rectiligne.*

*Les zones de type **« en grappe »** (depuis les années 70) se caractérisent par des rues sinueuses ou en impasse. Il existe plusieurs versions de zone en grappe : les zones en "raquette", en arborescence…*

*Dans les zones de type **« groupé »,** les maisons en rangée dominent.*
Remarque : Un autre type de zone existe, dit "ouvrier", qui n'a pas été pris en compte ici.

Figure C.9. *Une typologie des zones pavillonnaires en 4 classes illustrée à l'aide de données de la BD TOPO®*

Pour suivre les indicateurs dans le temps, la méthode d'observation n'impose pas de disposer d'un modèle de gestion spécifique des données temporelles. Ainsi, dans une exploitation élémentaire, nous montrons qu'il est possible d'effectuer un suivi d'évolution avec des données exploitant le modèle temporel le plus simple (le modèle par *snapshots* présenté dans le chapitre A.3), bien qu'il ne fournisse aucune information directe d'évolutions. Pour une exploitation plus développée, nous introduisons d'abord des fonctionnalités d'analyses géométriques, puis des données d'*appariement et de calculs différentiel*[9].

- Dans le premier cas, nous tentons d'identifier les créations de bâtiments, en utilisant seulement les fonctionnalités de superposition des SIG. Nous partons du principe que lorsqu'un bâtiment se crée en un lieu où il n'y avait pas de bâtiment auparavant, alors la composante géométrique de ce bâtiment ne se superposera pas avec la composante géométrique d'un autre bâtiment issu d'une version antérieure de la base (cf. Annexe 7). Les résultats obtenus sur la zone testée (près de Marne la Vallée) sont satisfaisants. Il ne peuvent cependant pas être généralisés aux zones densément urbanisées, car les constructions de nouveaux bâtiments correspondent alors plus souvent à des opérations de renouvellement urbain (destruction puis création) qu'à de véritables créations.

- Dans le second cas, nous utilisons des données issues de techniques d'appariement. Ces techniques qui restent en grande partie du domaine de la recherche ouvrent des perspectives concernant des observations de phénomènes plus dynamiques. Elles nous ont permis de ici de disposer de données d'évolutions détaillées (sur la nature de chaque changement de

9. Rappel : L'*appariement* est l'opération consistant à identifier dans deux bases de données différentes, deux représentations d'un même objet. Le *calcul différentiel* est l'opération qui consiste à identifier les variations entre deux représentations différentes.

données : création, destruction, mais aussi mutation géométrique et/ou sémantique).

En parallèle, nous exploitons le modèle temporel du MOS pour tester une méthode d'analyse statistique sur les mutations (ou couples d'états successifs), qui nous a permis ici de dégager des tendances générales sur l'évolution du phénomène d'urbanisation étudié. Cette méthode d'analyse pourrait être reproduite pour d'autres phénomènes recensés dans le MOS (ex. : suivi des grandes emprises d'activités, des surfaces commerciales).

C.2.2.3. Données utilisées, éléments techniques

Pour caractériser les 'zones pavillonnaires', nous avons utilisé deux bases de données différentes. La première, la BD TOPO® a fourni des données sur les 'bâtiments' et sur le 'réseau routier'. La base de données topographiques de l'IGN propose en effet une représentation du territoire dans une gamme d'échelle compatible avec le niveau d'observation élémentaire que nous souhaitions introduire. Par ailleurs, elle offre une représentation géométrique détaillée des 'bâtiments' et des 'tronçons de routes' qui permet d'effectuer de nombreuses analyses morphologiques. La BD TOPO® ne contenant pas d'objets 'zones pavillonnaires' (ou 'zone d'habitat individuel') en tant que tels, nous avons utilisé une deuxième source de données, la base MOS, pour l'exploitation élémentaire de la méthode.

Pour construire des caractéristiques d'observation, nous avons procédé de deux façons :

- dans un premier temps, nous nous sommes appuyés sur le MOS. Nous avons pu le faire car nous bénéficions de la *compatibilité temporelle* des bases de données exploitées : Les versions de la BD TOPO® sur la zone étudiée (une portion de territoire autour de Lagny près de Marne la Vallée) correspondent en effet aux dates de 1993 et de 1999 proches des dates de mise à jour du MOS (en 1994 et 1999).

- ensuite, dans la mise en œuvre élaborée, pour profiter d'une plus grande *compatibilité géométrique* entre les 'objets' observés[10], nous avons introduit les objets 'îlots routiers' obtenus par numérisation manuelle à partir de la BD TOPO®. En effet, la numérisation des 'îlots' s'appuyant sur les données du 'réseau routier' de la BD TOPO®, nous étions assurés de leur cohérence géométrique avec les autres 'objets' de la BD TOPO® ('tronçons de routes', 'bâtiments'). Nous avions par exemple la garantie d'échapper au cas où un bâtiment est « à cheval » sur deux zones pavillonnaires. Ces données 'îlots routiers' permettent aux utilisateurs de disposer d'une solution alternative lorsqu'il n'existe pas d'autres *données-support*.

10. La *compatibilité géométrique* des informations géographiques est le problème classique en *géomatique* présenté dans le chapitre B.2 : lorsqu'un utilisateur superpose dans un SIG des données issues de bases différentes, il est généralement confronté aux problèmes d'incohérences des supports géométriques ; Les tracés des limites et des axes des objets géographiques ne sont pas exactement identiques dans les deux bases.

Pour enrichir la méthode, nous avons profité de recherches menées au laboratoire COGIT : Cécile Duchêne a calculé pour nous la valeur de certaines caractéristiques géométriques (convexité du bâti, indicateurs sur la « rectilignité » des tronçons de route et sur le fait qu'ils soient des impasses) grâce aux outils d'analyses géométriques qu'elle a développés pour la *généralisation*. Arnaud Braun nous a fourni les données du différentiel entre les deux versions de la BD TOPO® utilisées en reprenant des travaux initiés par T. Badard. Pour cette partie, nous avons restreint la zone d'étude à une portion de territoire de la commune de Roissy en Brie regroupant une grande variété de zones pavillonnaires (cf. figure C.4).

C.2.3. Vers un outil d'observation et d'analyse de l'urbanisation de type habitat individuel

C.2.3.1. Enjeux d'une méthode de suivi de l'urbanisation pavillonnaire dans le domaine de l'aménagement du territoire

Les enjeux de cette application sont d'abord thématiques ; La mise en œuvre d'une observation de *l'urbanisation pavillonnaire* semble en effet répondre à un réel besoin des aménageurs, des urbanistes, des sociologues, des experts en transports [Cornaut, 2000], [Petit, 2000], [Millot, 2003]. Il s'agit ici de leur fournir des moyens pour observer les 'zones pavillonnaires' soit pour étudier le *phénomène d'urbanisation pavillonnaires* lui-même, soit pour étudier d'autres phénomènes ayant trait à ce type d'urbanisation. De l'avis même des thématiciens, il existe peu de données géographiques sur les 'zones pavillonnaires' en tant que telles. Celles qui existent permettent de dégager des tendances générales sur le phénomène (ex. : développement/régression, stagnation, localisation) et concernent uniquement une partie du territoire national. Or, au delà des tendances générales, on peut distinguer des formes variées d'urbanisation pavillonnaire et celles-ci influent sur de nombreux phénomènes étudiés en urbanisme (accidentologie, ségrégation sociale, etc.). Il importe donc de pouvoir saisir cette diversité. Les enjeux sont d'autant plus important que la portion de territoire concernée par ce phénomène est grande et qu'elle tend à s'accroître. Par ailleurs, dans le contexte actuel qui valorise les opérations de renouvellement urbain et condamne la périurbanisation comme consommatrice d'espace, la faible densité des zones pavillonnaires est de plus en plus mis en avant pour les critiquer. Les « zones pavillonnaires » sont ainsi l'objet de véritables débats en urbanisme. En proposant des moyens pour mieux les observer dans les SIG et en construisant une méthode pour les caractériser, nous offrons des outils pour mieux les étudier et mieux les connaître.

Les enjeux sont aussi techniques. Il s'agit en effet de développer des moyens d'observations pouvant s'adapter à des situations très variées (en fonction des données, outils et compétences disponibles, en fonction également du point de vue sur les objets *zones pavillonnaires*). La compatibilité de la solution proposée en particulier avec les données du MOS qui sont utilisées par plusieurs organismes centraux d'aménagement de la région Ile de France {Foulard, 2005}, pourrait également profiter aux utilisateurs travaillant sur l'un des quatre vingt trois autres 'objets' du MOS (par exemple pour le suivi de l'habitat collectif, des emprises d'activité, des jardins familiaux, …).

C.2.3.2. Enjeux de la méthode de suivi de l'urbanisation pavillonnaire pour les SIG

Cette application présente les enjeux généraux attachés à cette thèse, dans le cadre restreint de l'observation de l'urbanisation pavillonnaire. Elle illustre en particulier l'enjeu technique principal, qui se présente sous la forme d'un double défi, il s'agit de pouvoir :

- utiliser des SIG pour observer des objets sur lesquels on dispose de peu de données, et même éventuellement de s'adapter à des situations où il n'existe pas de données explicites sur ces 'objets' en tant que tels,

- suivre dans un SIG des évolutions sur ces objets, sans impliquer l'usage d'un modèle de données temporelles spécifique (cf. chapitre A.3).

Elle présente donc un enjeu stratégique pour les SIG : celui de proposer une solution pour observer des phénomènes géographiques, solution qui échappe à certaines critiques classiques formulées à l'égard des applications SIG concernant leurs coûts et leur complexité. En particulier, elle tente d'éviter la qualification d' « usine à gaz », (c'est à dire de solution souvent compliquée, manquant de visibilité de souplesse et coûteuse au regard des éléments fournis, ne rentabilisant pas les investissements consentis) {AUE, 2006}. Elle propose pour cela d'élaborer une *« solution adaptative »* où, par *adaptative*[11], nous signifions une solution qui :

- puisse intéresser des utilisateurs aux profils variés. En particulier, il faut qu'il existe une configuration simple adaptée aux utilisateurs disposant d'un minimum de moyens ;

- puisse évoluer en fonction des besoins, en s'enrichissant par exemple de compétences, données et outils acquis au fur et à mesure par les utilisateurs.

Cette application illustre également une évolution dans l'utilisation des SIG, où partant d'une exploitation comme outils d'observatoire, on aboutit à une utilisation comme outils d'analyse et d'aide à la décision. Elle montre en particulier que des caractéristiques (ex. : la densité) utilisées dans une observation statique (ex. : une carte de densité), si elles sont conçus comme des indicateurs de suivi dans une observation plus cinématique (ex. : étude de la densification), peuvent devenir des paramètres d'analyse dans une observation plus dynamique (ex. étude du lien entre la densification et la morphologie). Ce dernier niveau d'analyse où l'on cherche des corrélations et / ou des relations de causalité ouvre sur d'autres analyses proches de la modélisation et de la simulation urbaine.

11. Pour exprimer la capacité de leur logiciel à être une *solution adaptative* les fabricants d'ORACLE (système de gestion de bases de données) emploi le terme de *scalibility à la place de « l'adaptabilité»*.

C.3. Le phénomène d'imperméabilisation des sols

Le phénomène *d'imperméabilisation des sols* participe d'un phénomène plus général : le phénomène d'urbanisation. Les aménageurs s'y intéressent en particulier lorsqu'il s'agit de dimensionner les systèmes de drainages en prévision des risques d'inondation et de pollution des eaux pluviales (cf. figure C.10). Dans cette application nous l'appréhendons sous cet angle et nous proposons de le suivre à l'aide d'un indicateur principal : le *coefficient d'imperméabilisation*. Cet exemple d'application apporte un éclairage sur l'une des étapes principales de la méthode d'observation proposée : la construction des indicateurs de suivi. Il montre qu'il est possible d'effectuer des analyses qui ne se réduisent pas à des simples requêtes spatiales en exploitant les fonctionnalités des SIG et en intégrant des fonctionnalités externes (provenant par exemple d'outils statistiques). Dans le cadre de ces recherches, cette application forme avec la suivante deux illustrations contrastées prouvant que la méthode peut s'adapter à des profils d'utilisateurs différents. Pour un « suivi-bureautique » de phénomènes géographiques ou pour des utilisateurs non géomaticiens, on appliquera la méthode à partir d'outils standards. Pour des suivis plus techniques et/ou pour des utilisateurs plus expérimentés en SIG, l'exploitation fera intervenir des outils, des techniques plus sophistiquées portant sur des analyses d'informations sémantiques, topologiques et/ou géométriques.

http://www.febe-febelco.be/febelco/images/

Figure C.10. *Imperméabilisation des sols et systèmes de drainages*

Lorsque les sols sont revêtus de surfaces imperméables du fait des aménagements et constructions, il est indispensable de mettre en œuvre un système d'évacuation des eaux pluviales

http://pedagogie.ac-montpellier.fr/disciplines/

http://photos.linternaute.com/photo/593010/6509973407/193/intemperies_dans_le_gard/

Figure C.11. *Pollutions des eaux pluviales et inondations sont deux risques liés au ruissellement sur les zones imperméabilisées*

C.3.1. Suivi de l'imperméabilisation des sols pour réduire les risques liés au ruissellement

C.3.1.1. Introduction au coefficient d'imperméabilisation des sols

Avec l'urbanisation, les surfaces imperméables s'étendent. Les eaux pluviales qui ne peuvent pénétrer dans les sols ruissellent en augmentant les risques d'inondation et de pollution des eaux pluviales (fig. C.11). Les catastrophes qui ont eu lieu à Nîmes (1988) et dans d'autres villes ont fait prendre conscience de façon plus marquée de la nécessité d'une bonne gestion de ces risques. A l'occasion d'urbanisations nouvelles ou d'extensions d'urbanisations existantes, il importe de pouvoir suivre et de prévoir les surfaces imperméabilisées pour dimensionner au mieux les

canalisations et autres ouvrages hydrauliques. En fait, l'aménagement hydraulique d'une zone nécessite de prendre en compte de nombreux éléments : sa topographie naturelle (plus le terrain est pentu moins l'eau a de temps pour pénétrer dans les sols), la pluviométrie (plus le volume d'eau tombée est important, plus il faut de capacité pour l'évacuer), la nature des sols (l'eau pénètre plus ou moins facilement dans le sol suivant sa nature), etc. Tous ces indicateurs spatiaux peuvent être introduits sous forme d'attributs dans une base d'occupation des sols vecteur ou sous forme de champs raster, pour aider à spécifier les réseaux techniques d'évacuation des eaux pluviales. Dans la pratique, on utilise le plus souvent un unique paramètre théorique qui intègre ces différents facteurs : le *coefficient de ruissellement Cr*[12]. Cependant, sous certaines conditions, Caquot montre qu'il est possible de lui substituer le *coefficient d'imperméabilisation C* plus simple à établir[13] [Chocat, 1997]. Pour que *C* soit une bonne approximation de *Cr,* il faut que:

- l'on se trouve en milieu urbain, c'est à dire que les surfaces perméables ne représentent pas plus de 20% de la surface totale et que

- la pente du terrain ne soit pas trop importante : elle doit être inférieure à 5%. Sinon, les surfaces perméables se comportent quasiment comme des surfaces imperméables : la pluie n'a pas le temps de pénétrer dans le sol, il devient alors nécessaire de prendre en compte leur contribution au ruissellement.

Cette approximation (*Cr≈C)* revient à ne considérer que la contribution des zones imperméabilisées au ruissellement et à supposer que cette contribution s'effectue avec des pertes négligeables[14]. En se plaçant dans les conditions de validité du modèle de Caquot, une méthode de suivi du *coefficient d'imperméabilisation* dans un SIG constituerait alors un outil d'aide à la gestion, voire d'aide à la décision, pour les hydrologues et les aménageurs.

C.3.1.2. Observation de l'imperméabilisation des sols

L'observation de *l'imperméabilisation des sols* dans un SIG ne soulève a priori pas de problème conceptuel. En effet, il est aisé de modéliser les *zones imperméabilisées* en information géographique et de construire des données pour les représenter : leur emprise sur le terrain s'identifie facilement à partir d'une photographie aérienne, ce qui permet de les obtenir sous forme de données géographiques par simple numérisation (cf. figure C.12). De plus visuellement, il est simple de suivre leurs évolutions, en comparant deux états différents (cf. figure C.13). La photo-interprétation suivie du « décalquage » des surfaces imperméabilisées est d'ailleurs la démarche traditionnelle pour calculer et mettre à jour la valeur des surfaces imperméabilisées dans les études hydrologiques [Deutsch, 1989]. Toutefois, si cette application ne pose pas de difficulté théorique, sa mise en œuvre pratique reste laborieuse, même lorsque le « décalquage » (papier) est remplacé par la digitalisation (numérique). Ainsi, même si la forme numérique permet dans un SIG de calculer immédiatement la superficie de la surface imperméabilisée, le travail préalable de numérisation de chaque zone demeure un travail fastidieux, qu'il faut reproduire à chaque mise à jour… Par ailleurs, on note ici que, relativement à la gestion des systèmes de drainage, ce qui intéresse les hydrologues ce n'est pas tant la *zone imperméabilisée* elle-même, mais sa *superficie*, car à travers elle, les aménageurs voient le moyen d'obtenir facilement le *coefficient d'imperméabilisation* (par simple division par la superficie totale). Ainsi, une méthode de suivi de l'imperméabilisation des sols, à partir d'une détermination directe et automatique du *coefficient d'imperméabilisation* pourrait ici leur convenir aussi bien.

12. Coefficient de ruissellement

$$C_r = \frac{P_n}{P_b}$$

P_n est la quantité de pluie nette, c'est à dire la quantité d'eau de pluie tombée qui participe du ruissellement sur un site A fixé
P_b est la quantité totale d'eau tombée sur la surface de A.

13. Coefficient d'imperméabilisation

$$C = \frac{S_{imp}}{S_{totale}}$$

S_{imp} est la surface imperméabilisée sur un site A donné
S_{totale} est la surface totale du site A

14. En effet, si on assimile Cr à C, on a alors l'égalité :

$$\frac{P_n}{P_b} = C_r \approx C = \frac{S_{imp}}{S_{totale}}$$

En introduisant h, la hauteur d'eau moyenne tombée sur la zone étudiée, alors par définition on a l'égalité :

$$P_b = h * S_{totale}$$

entre le produit de h par la surface totale et la quantité d'eau totale tombée sur la zone. D'où, par identité l'approximation de P_n :

$$\frac{P_n}{P_b} \approx \frac{h * S_{imp}}{h * S_{totale}} \Rightarrow P_n \approx h * S_{imp}$$

Ce qui revient à dire que la quantité d'eau qui ruisselle, correspond à la totalité h d'eau tombée (en négligeant les pertes), uniquement sur les zones imperméabilisées.

a.

b.

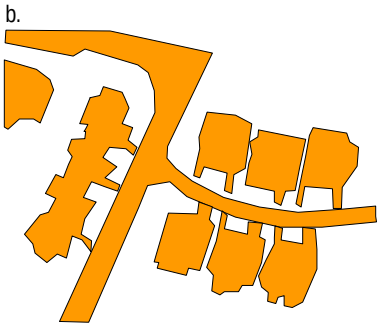


Figure C.12. *A partir d'une photographie aérienne (a), on peut numériser des données pour représenter les zones imperméabilisées (b)*

Figure A.13. *Observation visuelle des changements (en rouge) et du maintien (en jaune) des zones imperméabilisées*

C.3.2. Vers un suivi du coefficient d'imperméabilisation

Dans cette application, en collaboration avec Jean-Claude Deutsch du laboratoire CEREVE[15] et avec l'apport de Marie Mayeux et d'Olivier Bonin, nous proposons de mettre au point et de tester les principes d'une méthode alternative pour suivre l'évolution du *coefficient d'imperméabilisation* [Bordin et al, 2005]. Lorsqu'il n'existe pas de données, plutôt que de numériser les 'zones imperméabilisées', nous proposons d'évaluer la superficie de ces zones. Pour cela nous partons de deux postulats :

- chaque objet géographique génère une certaine surface imperméable dans son voisinage immédiat dépendant de son type et de sa fonction (fig. C.14),

- deux objets géographiques de même type, ayant la même fonction, (ce que nous traduirons en disant qu'ils appartiennent à une même *catégorie*) génèrent statistiquement la même surface imperméabilisée.

Le principe de la méthode est d'établir préalablement à l'observation, une classification des 'objets topographiques' (micro) en *catégories* C_i, telle que pour chaque C_i, on puisse définir une relation statistique R_i telle que R_i permette de calculer une estimation $\acute{s}_{ij}$ de la surface imperméable élémentaire s_{ij} associée réellement à chaque objet o_{ij} de la catégorie C_i, soit :

$$\forall\, C_i, \exists\, R_i \;/\; \forall o_{ij} \in C_i,\ R(o_{ij}) = \acute{s}_{ij} \text{ où } \acute{s}_{ij} \approx s_{ij}$$

Disposant d'une relation R_i pour tous les types d''objets topographiques' $\{o_{ij}\}_{\forall ij}$, il s'agira alors de définir une partition de l'espace en 'objets méso-structurel' pour calculer une approximation $\acute{S}_{imp}$ de la surface imperméabilisée S_{imp} associée à ces 'portions de territoire'. Cette estimation sera obtenue en agrégeant les 'objets topographiques' selon l'emprise de chaque 'objet méso-structurel' et en calculant à l'aide des R_i préalablement définis l'indicateur :

somme des estimations $\{\acute{s}_{ij}\}_{\forall ij}$ des surfaces imperméabilisées élémentaires associées à chaque 'objet topographique' participant à l'agrégation

$$\acute{S}_{imp} \approx \Sigma\, \acute{s}_{ij}\ (\forall o_{ij} \in C_i,\ \forall C_i) / o_{ij} \subset \text{emprise de l''objet méso'}$$

Une estimation C du *coefficient d'imperméabilisation* sera alors calculée pour caractériser chaque 'objet méso' définie par :

$$C = \acute{S}_{imp} / S_{(\text{de l'emprise d'agrégation})}$$

Figure C.15. *Principe de la méthode de suivi du coefficient d'imperméabilisation par estimation de la superficie de la surface imperméabilisée*

15. CEREVE : Centre d'Enseignement et de Recherche Eau Ville Environnement à l'ENPC (Ecole des Ponts et Chaussées)

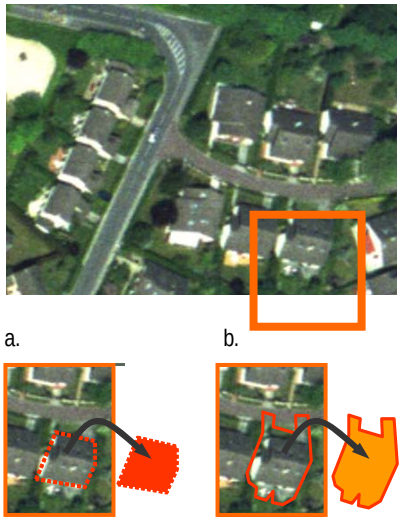


Figure C.14. *Exemple de surface imperméable associée à un pavillon*

Sur la photographe aérienne, on distingue l'emprise du pavillon représenté dans la figure a. de l'emprise de la zone imperméable associée au pavillon représentée dans la figure b (qui peut comprendre l'allée centrale, l'accès au parking, une ou deux terrasses, etc.) . Numérisées sous forme de données vecteurs, elles correspondent à des objets distincts. Dans les base de donnés, nous disposons de l'emprise du bâti, mais pas de l'emprise de la surface imperméabilisée.

Cette méthode n'implique pas de reconstituer la composante géométrique des 'zones imperméables' pour obtenir leur superficie, puisqu'elle est obtenue par calcul à partir de relations statistiques. De fait, nous n'observons pas les 'zones imperméables' elles-mêmes. En revanche, l'utilisation des outils SIG reste requise :

- dans l'élaboration des relations statistiques : les fonctionnalités des SIG servent à extraire commodément les différentes caractéristiques des objets élémentaires (géométriques et sémantiques) pour tester leur pertinence en tant que paramètre dans les relations R_i. Par exemple, pour les objets de la catégorie *route*, la caractéristique *classement fonctionnel* est un paramètre significatif pour la relation R_i, ce qui n'est pas le cas du *sens de circulation*.

- dans l'exploitation ultérieure de la méthode : Pour reconstituer une approximation de la superficie de la surface imperméable totale sur une zone d'étude donnée, les fonctionnalités du SIG serviront à sélectionner les objets présents dans la zone étudiée et à identifier leurs caractéristiques.

C.3.3. Mise en oeuvre de l'observation de l'imperméabilisation des sols

Cette application consiste essentiellement à mettre au point et à tester une méthode pour construire des relations R_i, liant l'emprise d'un 'objet topographique' o_{ij} de catégorie i, (c'est à dire ayant un type et une fonction donnés) à la surface imperméabilisée qui lui est associée (cf. figure C.15).

Dans la première étape, nous choisissons une catégorie C d''objets élémentaires' (par exemple le 'bâti'). Un jeu de données test est ensuite constitué en choisissant un ensemble $\{o_j\}$ d''objets' de C et en numérisant les surfaces $\{s_j\}$ associées à chaque o_j (fig. C.4.16a). L'utilisation de la même méthode que celle traditionnellement utilisée pour obtenir la valeur des surfaces imperméables, nous permet de considérer les valeurs des superficies s_j comme des valeurs vraies et de disposer ainsi d'un jeu de surface de référence.

La seconde étape consiste à établir une relation statistique R permettant d'estimer la superficie des 'surfaces' s_j en fonction des caractéristiques des 'objets' o_j, (en particulier en fonction de caractéristiques de leur composante géométrique : la superficie pour les objets polygonaux, la longueur pour les objets linéaires) (fig. C.4.16b). La validité de la relation ainsi obtenue doit ensuite être testée sur d'autres jeux de données.

Enfin, en passant en revue l'ensemble des 'objets topographiques' et en les classant dans des catégories pertinentes, en établissant ensuite les différentes formes de relations R_i liant les objets d'une catégorie C_i à la surface imperméabilisée qu'ils génèrent, on peut alors espérer calculer automatiquement le coefficient d'imperméabilisation associé à un territoire. On peut également envisager de prévoir le coefficient d'imperméabilisation à partir de la planification de l'urbanisation.

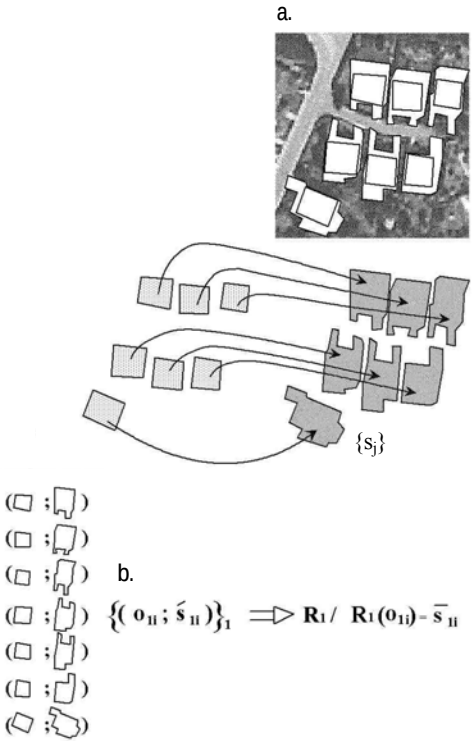


Figure C.16. *Les étapes principales de la méthode de construction des R_i*

Dans un premier temps on élabore un jeu de données d'essai, en sélectionnant des objets et en numérisant leurs surfaces imperméabilisées associées (a). Dans un second temps, ayant les couples (oj ; sj), on cherche à établir une relation statistique R permettant de déduire une estimation $\hat{s}_j$ de s_j à partir des caractéristiques de o_i (b). La troisième étape consiste à valider la relation R trouvée à partir d'un autre jeu de données.

C.3.3.1. Données utilisées, éléments techniques

Pour disposer de données géographiques sur des objets élémentaires, nous proposons d'utiliser les données de la BD TOPO®. Cette base possède en effet plusieurs qualités importantes pour cette application :

- elle comporte des objets élémentaires pertinents avec le niveau d'observation des zones imperméables (tels que le bâti et les routes).

- ses spécifications sont homogènes sur tout le territoire, ce qui assure la *faisabilité* de la méthode en dehors de la zone de test,

- la pérennité de ses spécifications dans le temps est prévue, ce qui garantit de pouvoir *comparer* les observations effectuées à des dates différentes.

Pour cette étude, nous avons travaillé sur la région de Marne-la-Vallée. Les jeux d'essais ont été constitués à partir des versions de 1994 et 1999 de la BD TOPO® et de numérisations de photographies aériennes de l'IGN de 1993 et de 1999.

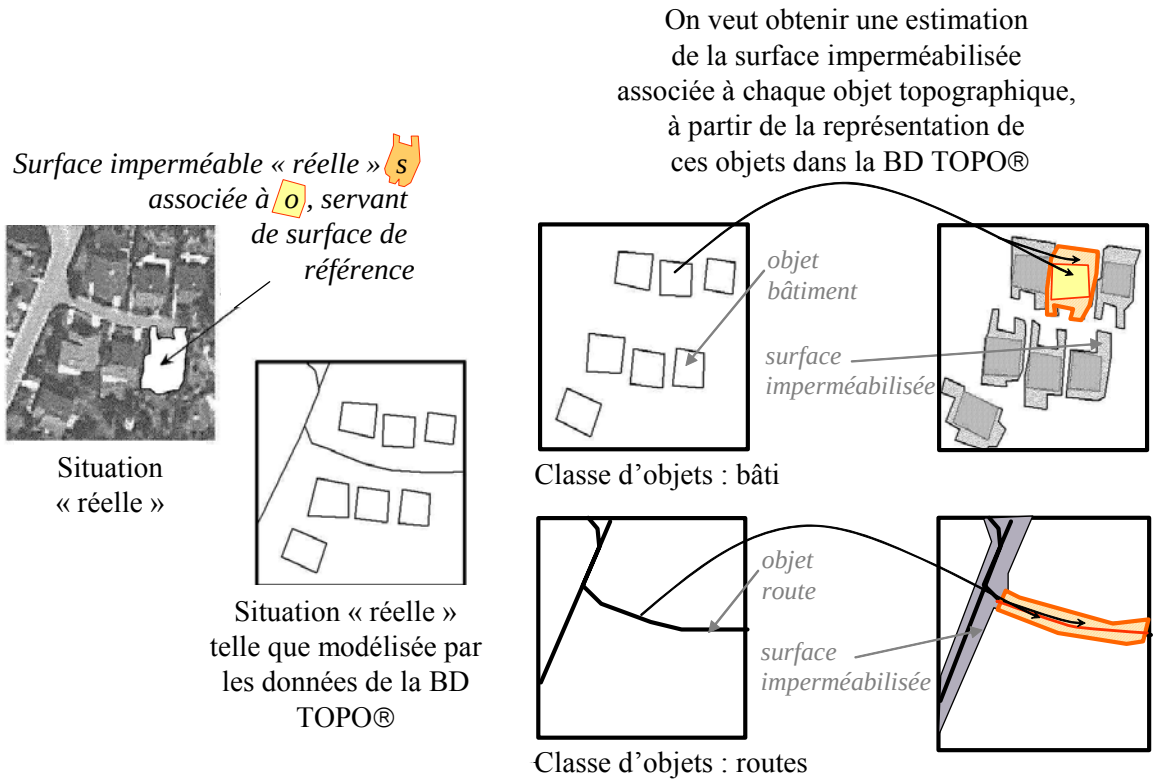

On veut obtenir une estimation de la surface imperméabilisée associée à chaque objet topographique, à partir de la représentation de ces objets dans la BD TOPO®
Surface imperméable « réelle » s associée à o, servant de surface de référence
Situation « réelle »
Situation « réelle » telle que modélisée par les données de la BD TOPO®
objet bâtiment
surface imperméabilisée
Classe d'objets : bâti
objet route
surface imperméabilisée
Classe d'objets : routes

Figure C.17. *Principe de la méthode de suivi du coefficient d'imperméabilisation à l'aide de la BD TOPO®*

Partant de la situation « réelle » considérée au travers des photographies aériennes, nous avons exploité la modélisation de la réalité donnée par la BD TOPO® (cf. figure C.17). Le document sur les spécifications de la BD TOPO® détaille cette modélisation en précisant les objets géographiques qui y sont représentés ('réseaux routiers', 'bâti', 'réseaux hydrographiques', 'voies ferrées', 'limites administratives', etc.). Les 'zones imperméables' n'en font pas partie. Pour tester le principe et la méthode d'estimation des surfaces imperméabilisées élémentaires, nous avons retenus deux grands *thèmes* d'objets topographiques : les *routes* et le *bâti*. Pour chacun nous avons repris des catégories proposées par la BD TOPO® :
- Pour les 'routes', les catégories correspondent au classement : « autoroute », « nationale », « départementale », « autre route »,
- Pour les 'bâtiments', les catégories sont : « Bâtiment-industriel-agricole », « Bâtiment-quelconque »

Les résultats obtenus sont particulièrement intéressants pour les 'routes' *:* les relations R_i trouvées donnent de bonnes approximations des surfaces imperméabilisées générées par les différents types de voie. Pour les 'bâtiments', les résultats sont moins satisfaisants. Il faudrait affiner les *catégories*. En effet comme l'illustre la figure C.18, dans la BD TOPO®, deux 'bâtiments quelconques', ayant une emprise au sol identique, peuvent correspondre à des réalités très différentes et donc être associés à des surfaces imperméables également distinctes. La composante géométrique de certains pavillons peut se confondre avec celle d'une petit collectif (de

même celle de certains petits collectifs avec des immeubles plus importants). Or, dans les deux cas les besoins en infrastructures annexes (de circulation, de stationnement, de jeux, de jardins, etc.) ne sont pas les mêmes. Il en résulte une imperméabilisation des sols différente. Une sous catégorisation de l'habitat (correspondant aux 'bâtiments-quelconques') en 'bâtiments individuels' et en 'bâtiments collectifs' pourrait permettre d'améliorer ces résultats. En particulier, l'identification des 'zones pavillonnaires' servirait à différencier cette catégorie particulière d'habitat et permettrait d'affiner le calcul du coefficient d'imperméabilisation sur ces zones. Dans ces travaux nous avons testé le principe et la méthode de calcul. Pour une application opérationnelle du calcul du coefficient d'imperméabilisation, il faut développer ces premiers éléments et prendre en compte l'ensemble des objets de la BD Topo®. Il faut pour cela prévoir l'apport d'une expertise thématique (en architecture, en urbanisme, en aménagement) pour construire des catégories plus pertinentes, pour aboutir à des relation R_i plus précises.

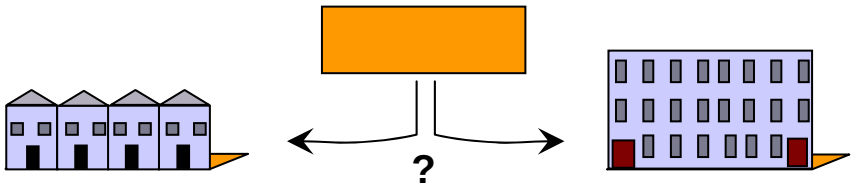


Figure C.18. *Une « même » donnée* bâtiment quelconque *peut correspondre à des réalités très différentes*

C.3.4. Vers un outil opérationnel d'observation de l'imperméabilisation des sols au service des aménageurs

C.3.4.1. Enjeux d'une méthode de suivi du coefficient d'imperméabilisation dans le domaine de l'aménagement du territoire

L'objectif opérationnel de cette application est le développement d'un outil d'aide à l'aménagement du territoire et plus précisément de gestion des aménagements hydrauliques. Afin d'y répondre nous proposons une méthode pour établir des relations entre le *coefficient d'imperméabilisation* et des paramètres urbains, de façon à calculer ce coefficient pour l'urbanisation actuelle et pour simuler sa valeur dans une urbanisation future ou projetée. La mise au point d'une telle méthode de suivi automatique du *coefficient d'imperméabilisation* représente un double enjeu pour les thématiciens :

- un enjeu technique. La méthode traditionnelle généralement utilisée pour calculer et mettre à jour la valeur des surfaces imperméabilisées dans les études hydrauliques est longue et donc coûteuse. Construire une méthode qui permette aux utilisateurs d'éviter la numérisation de ces 'zones' tout en produisant l'information dont ils ont besoin : la *superficie de la surface* imperméabilisée à jour (son emprise et sa localisation ne sont pas nécessaires) contribuerait à faciliter la gestion et la décision dans les études d'aménagements hydrauliques.

- un enjeu stratégique. Le *coefficient d'imperméabilisation* intervient à l'origine d'une chaîne d'études techniques et thématiques concernant des phénomènes de plus en plus globaux : il est un paramètre essentiel pour dimensionner les systèmes de drainages, ces systèmes sont eux-mêmes des éléments fondamentaux pour la gestion du *ruissellement*, phénomène partiellement responsable des *inondations* et des *pollutions*… Aider à

déterminer ce coefficient revient ainsi à participer modestement à l'ensemble constitué par ces études.

La méthode d'observation du coefficient d'imperméabilisation proposée en établissant des relations R_i peut également contribuer à des études stratégiques concernant des projets d'aménagements du territoire : En se plaçant dans une perspective prospective, ces relations peuvent servir à tester différentes options d'aménagements en simulant facilement les surfaces imperméables générées dans chaque cas. Au-delà d'une méthode d'observation de l'imperméabilisation des sols, ces travaux pourraient également être considérés comme une contribution à l'observation de *l'urbanisation*.

C.3.4.2. Enjeux d'une méthode de suivi du coefficient d'imperméabilisation avec un SIG

Le suivi des surfaces imperméabilisées est un enjeu suffisamment important pour que l'on ait déjà tenté d'utiliser les SIG pour l'améliorer [environnement-transport.hauts-de-seine.net]. Cependant, les expériences précédentes ont exploitées des bases de données d'occupations des sols qui se sont avérées trop générales pour permettre d'obtenir de bonnes estimations ; en effet deux portions de territoire ayant la même 'occupation des sols', peuvent être associées à des surfaces imperméabilisées très différentes (comme nous avons pu le constater sur la seule occupation : *bâti*). Jusqu'à présent, il n'a donc pas été possible d'effectuer le suivi du phénomène *d'imperméabilisation des sols* à partir de données vecteurs dans un SIG. D'autres recherches exploitent des représentations raster-images et s'appuient sur des analyses et traitements d'images. Les résultats obtenus semblent meilleurs. Cependant, dans ces travaux la modélisation des *objets géographiques* repose sur un ensemble de pixels homogènes, et n'existe pas sous la forme d'*objets informatiques* (cf. paragraphe A.1.2.1.1). Dès lors il devient difficile l'utiliser ces méthodes pour travailler en mode prospectif et tester des options d'urbanisation en fonction du nombre et des types d'aménagement envisagés. L'enjeu de cette application est donc de fournir à la fois un *outil de gestion* commode pour élaborer le coefficient technique d'imperméabilisation des sols, et un *outils d'aide à la planification* d'aménagements.

C.4. Le phénomène d'ambiance urbaine

L'observation de *l'ambiance urbaine* dans un SIG pose de façon conceptuelle le problème de l'*objet* principal d'observation. En effet, si les *zones pavillonnaires* ou les *surfaces imperméables* s'imposent comme des objets « naturels » pour suivre les effets de *l'urbanisation pavillonnaire* et de *l'imperméabilisation des sols*, il est plus difficile de désigner un objet géographique qui puisse être considéré comme le produit du phénomène *d'ambiance urbaine* et dont le suivi puisse servir au suivi du phénomène lui-même ; les effets du phénomène de l'ambiance urbaine ne produisent pas d'objets géographiques physiques. Ses effets sont essentiellement intangibles. Or pour suivre l'ambiance urbaine dans un SIG, il est indispensable d'avoir identifié un objet géographique, dont l'observation permette de tirer des éléments pour la décrire.

Cette application met en lumière le travail préalable à toute utilisation des outils SIG qui consiste à déterminer des éléments caractéristiques à observer, à les spécifier et à les formaliser en informations géographiques ou en indicateurs, puis en données géographiques et attributs. Cette étape préliminaire de modélisation et de spécification n'est pas particulière à la méthode d'observation que nous proposons. Cependant, elle y joue un rôle fondamental et central :

– pour spécifier l'objet principal d'observation servant au suivi du phénomène étudié, en particulier lorsqu'aucun 'objet' ne s'impose naturellement,

– pour trouver des éléments caractéristiques à partir desquels construire des indicateurs de suivi.

Pour s'assurer que les données informatiques produites conviennent à l'étude du phénomène considéré, il importe que ce travail soit mené sous forme de collaboration entre thématiciens et géomaticiens. Nous avons ainsi travaillé avec l'architecte et urbaniste, Maria Basile, qui s'intéresse en particulier au suivi de l'ambiance urbaine, dans le cadre de ses travaux sur l'impact de la « gestion de centre ville » sur les transformations urbaines [Basile, 2005].

C.4.1. Définition de la notion d'ambiance, différentes approches

C.4.1.1. Un phénomène difficile à représenter dans un SIG

Le phénomène de *l'ambiance urbaine* est difficile à saisir, en partie car il est immatériel. Il ne s'inscrit pas physiquement dans le territoire : Contrairement aux *zones pavillonnaires* identifiables sur le terrain (fig. C.19), l'*ambiance* en tant que fait, ne peut être digitalisé à partir d'une photographie aérienne. Certains « objets géographiques » immatériels existent, qui sont malgré tout appréhendés dans les SIG, par exemple, une *zone de chalandise*. Leur emprise n'est pas marquée sur le terrain, mais elle est donnée comme une fonction d'un certain nombre de paramètres. L'*ambiance urbaine* est également fonction de différents facteurs, mais ils n'appartiennent pas tous au même registre : *l'ambiance* est multi-dimensionnelle. L'enseignement en « maîtrise des ambiances » dans les écoles d'architectures l'illustre bien, qui propose des présentations dans différentes matières telles que l'acoustique, la thermique et l'éclairement, abordant chacune une dimension de l'ambiance. De cette pratique résulte une multitude de représentations d' « ambiances », qui correspondent à

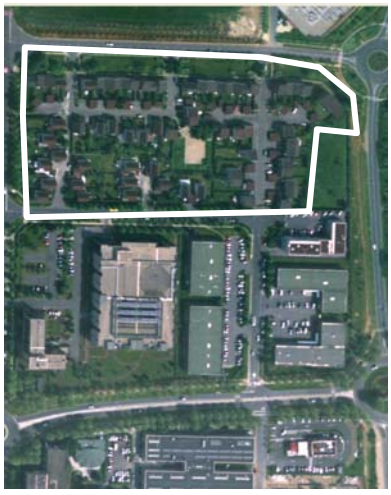


Figure C.19. *Inscription dans le territoire des objets zones pavillonnaires*

autant de façons de l'aborder. L'*ambiance* peut cependant être considérée comme un tout unique : *« une atmosphère matérielle et morale qui environne un lieu, une personne »*[16]. Toutefois, la difficulté principale pour saisir l'*ambiance urbaine* dans un SIG, vient du fait que ce qui la définit comme un tout varie d'un individu à l'autre et même d'un moment à l'autre, en fonction de l'humeur de celui qui la perçoit. Dans ces conditions il est difficile de la spécifier pour la modéliser sous forme de donnée géographique.

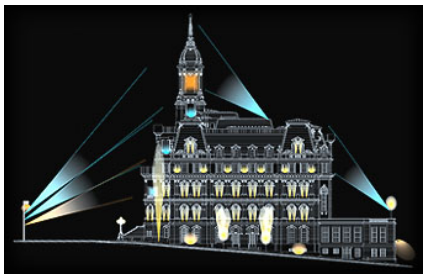
16. Dans l'ouvrage collectif du Laboratoire CRESSON [Amphoux, Thibaud, Chelkoff, 2004, p.18] Jean François Augoyard présente cette définition comme l'une des plus utilisées.

17. Cité par [Basile, 2006].

C.4.1.2. Deux approches principales pour observer l'ambiance : physiologique et cognitive

Le terme d'ambiance est apparu dès 1885. Il se définissait alors comme : *« environnement, atmosphère qui enveloppe une personne ou une chose*[17] *»*. Cette définition reste actuelle. On tend seulement aujourd'hui à préciser si l'environnement est considéré comme physique et/ou psychologique. Cette double facette de l'ambiance se retrouve d'ailleurs dans la définition de l'adjectif dérivé « ambiant », donnée par le Petit Robert. En effet, les exemples utilisés pour en illustrer le sens se réfèrent à des aspects techniques (« air ambiant, température ambiante ») et la citation renvoie quant à elle à une dimension plus psychologique (« soulevé par cet enthousiasme ambiant… » citation de Daudet).

Les deux facettes de l'ambiance sont souvent indissociablement liées, dans la perception globalisante que l'on en a. La gestion de l'éclairage des espaces publiques l'illustre bien (fig. C.20). Outre sa dimension fonctionnelle, il lui *« est aussi reconnue une utilité en terme de sécurisation des personnes et des biens et en terme d'embellissement de la ville : c'est dans ces aspects qu'intervient la notion d'ambiance. De nouvelles démarches visent à améliorer la vie sociale par l'intervention dans le cadre dans lequel elle se déroule : « les plans-lumière » structurent la ville pour que les éclairages ne se limitent pas à garantir une diffusion homogène de la lumière, mais aussi pour mettre en valeur les éléments émergeants du tissu urbain. Cette « mise en scène » des espaces urbains est accentuée avec l'apparition de nouveaux métiers, tels que celui de « concepteur lumière » »* [Basile, 2005, p.5]. Les collectivités qui tentent d'optimiser la gestion des candélabres utilisent parfois une application-SIG dédiée, développée par les constructeurs de logiciels, celle-ci fait souvent partie du « package » des applications SIG classiques proposées aux collectivités (à coté de la gestion du cadastre, de la voirie, des secteurs scolaires, etc.).



http://www.vieux.montreal.qc.ca/planlum/eng/tempsa.htm

Figure C.20. *Projet d'éclairage public*

Des travaux divers sont conduits pour appréhender *l'ambiance*. Certains privilégient l'approche technique et l'abordent au travers de mesures instrumentales (ex. : nombre de lux, de décibels). Elles aboutissent à des techniques de simulations sophistiquées intégrant de nombreux paramètres et peuvent produire des représentations très proches du réel. Par exemple, le logiciel SOLENE, développé au sein du CERMA, intègre différentes variables (trajectoire apparente du soleil, éclairage directe et indirect, etc.) pour proposer une promenade virtuelle à différentes heures de la journée et dans des périodes de l'année variées. Cette approche principalement mathématique ne présente qu'une partie de l'ambiance. Elle restitue une représentation de l'environnement physique partagé par tous. D'autres approches, du registre des sciences humaines (par exemple sur la dimension sociale des comportements dans l'espace ou sur les théories de la psychologie de la perception), l'appréhendent de façon plus globale. Elles considèrent des caractéristiques plus qualitatives (on jugera de la convivialité, de l'animation, de la sécurité d'un quartier ou d'une rue)

résultant à la fois de perceptions psychologiques (ex. : bien-être, peur), de perceptions physiologiques (perceptions olfactives, visuelles ou sonores) ou même de facteurs quantitatifs (ex. : nombre de passants par heure). On aboutit alors à une représentation de l'ambiance plus qualitative mais aussi plus relative. Ces travaux utilisent des méthodes et des outils proches de ceux de la psychologie cognitive. Il s'agit d'enquêtes-terrain, d'interviews, de questionnaires et d'analyses de texte, etc.

Enfin, on peut citer les travaux en géomarketing, même si ces recherches sont plus orientées vers la compréhension du fonctionnement des *aires de chalandises* que vers l'étude explicite de l'*ambiance urbaine*. Dans une approche plus économique qu'environnementale, elles visent à identifier les paramètres qui influent sur l'attractivité d'une zone pour essayer d'adapter l'environnement à la clientèle ciblée… Ces travaux établissent cependant une passerelle entre les deux points de vue sur *l'ambiance*: le point de vue technique, partiel mais collectif et le point de vue psychologique, global mais relatif. En effet, pour discriminer certains groupes de personnes, les « géomarketeurs » tentent d'adapter au mieux l' *« expérience »* d'une portion de territoire (notion proche de *l'ambiance* mais considérée d'un point de vue cognitif individuel) aux attentes des clients qu'ils cherchent à attirer, et pour cela ils étudient l'impact des perceptions sensorielles sur les perceptions psychologiques. Les représentations de '*l'ambiance urbaine'* en géomarketing intègrent ainsi une large gamme de perceptions, mais sa caractérisation est guidée principalement par l'aspect économique. L'ambiance est ainsi essentiellement appréhendée au travers des commerces et de leurs activités. Ces travaux utilisent des SIG pour analyser l'ambiance urbaine, mais l'ambiance observée, bien qu'à la fois technique et cognitive, ne correspond pas exactement à celle qui intéresse les collectivités. Sa thématique est trop spécialisée. Les commerces participent certes des centres villes, mais les aménageurs se penchent également sur des thèmes plus urbanistiques ou sociaux … Avec cette application, nous avons voulu contribuer à l'observation d'une ambiance plus globale et moins sectorielle.

C.4.2. Vers une observation globale de l'ambiance urbaine

C.4.2.1. Quel objet géographique pour observer l'ambiance ?

L'objet de cette application consiste à formaliser le suivi d'une ambiance urbaine qui ne serait pas uniquement commerciale, technique ou « ressentie ». *« Il nous a paru intéressant de dépasser les clivages entre disciplines et de tenter d'aborder l'Ambiance de façon plus globale, au plus près de ce qui est perçu au quotidien. Il s'agissait pour nous d'identifier des éléments participant de l'Ambiance en tant que tout, au-delà d'un simple processus d'addition. Sans prétendre saisir l'Ambiance en elle-même, nous avons cherché à définir des indices objectivables et descriptibles [et pouvant être suivis]. Le SIG, en tant qu'outil d'analyses localisées pluridisciplinaires, nous a semblé un support particulièrement pertinent pour concevoir de tels indicateurs»* [Basile et Bordin, 2005]. Il s'agit donc pour nous d'identifier de tels indices et de les modéliser dans un SIG sous forme d'indicateur, c'est à dire comme attribut d'un objet considéré comme caractéristique de l'ambiance et pouvant servir au suivi de cette ambiance.

Pour pouvoir « observer l'ambiance » dans un SIG avec la méthode d'observation multi-niveaux, il fallait définir un 'objet géographique' méso. Nous inspirant de l'étymologie du mot ambiance, dérivant du latin « ambire » signifiant faire le tour, nous avons cherché un *objet*

géographique qui puisse constituer un environnement à l'échelle des perceptions du passant. Nous nous sommes alors intéressées à la *rue*. En fait, nous avons considéré un objet proche du *tronçon de rue,* qui puisse servir à regrouper des *objets élémentaires* pertinents pour l'étude du phénomène. Plusieurs *objets élémentaires* ont été envisagés (les *espaces ouverts*, le *mobilier urbain*, les *bâtiments*, etc.). Pour une observation dans un SIG, le *bâti* nous a d'abord paru intéressant. En effet, les 'bâtiments' sont des objets géographiques qui servent de support à de nombreuses autres informations. 'Objets' généralistes, ils servent à localiser des informations descriptives données à l'adresse, par le biais du *géocodage*[18] Par ailleurs les 'bâtiments' interviennent dans de nombreuses thématiques ; on peut ainsi les caractériser d'un point de vue fonctionnel, architectural, social, etc. Dans la BD TOPO® que nous avons utilisée, l'information 'bâtiment' est disponible, mais elle comporte peu d'informations attributaires pour le caractériser. Pour compléter ces informations, nous avons tenté de relever sur le terrain des caractéristiques architecturales (en particulier sur leur état de conservation, le niveau d'insalubrité, leur style, etc.). Cependant, nous ne sommes pas parvenues à définir des spécifications suffisamment pertinentes et discriminantes. De plus en centre ville, la contiguïté du bâti, ne permet pas de dissocier les immeubles voisins (cf. figure C.21). C'est pourquoi nous avons finalement renoncé à caractériser l'ambiance urbaine à partir des 'bâtiments', pour nous concentrer sur d'autres aspects. Nous avons choisi les *commerces* comme objets d'études et donc comme 'objets micro', en considérant à la fois le contenant (c'est à dire le local, la *boutique*, dans une conception proche du bâti initialement envisagé) et le contenu (c'est à dire l'*activité* commerciale qui s'y exerce). Bien que chaque *boutique* participe à l'ambiance, elle ne permet pas individuellement de la représenter. Par contre, en application de la théorie du *Gestalt*, on sait que l'agrégation des éléments *commerces* contribue à *l'ambiance* en tant que perception globale. Ainsi, les commerces s'avéraient des 'objets micro' thématiques pertinents pour observer le phénomène d'ambiance urbaine. Il ne nous manquait plus qu'à définir la 'portion de territoire' support. Ne disposant pas d''objet méso' naturel, ni même d''objet géographique' pour servir d''objet méso-structurel' nous avons été conduit à construire un objet spécifique (cas c1, dans la définition de l'objet méso-structurel, cf. p. 143).

18. Connaissant l'adresse de chaque bâtiment, on peut alors placer sur la carte (ou intégrer dans la base de données) toute information ou caractéristique décrite par une adresse : il suffit de l'associer au bâtiment portant la même adresse. C'est le principe du *géocodage*.

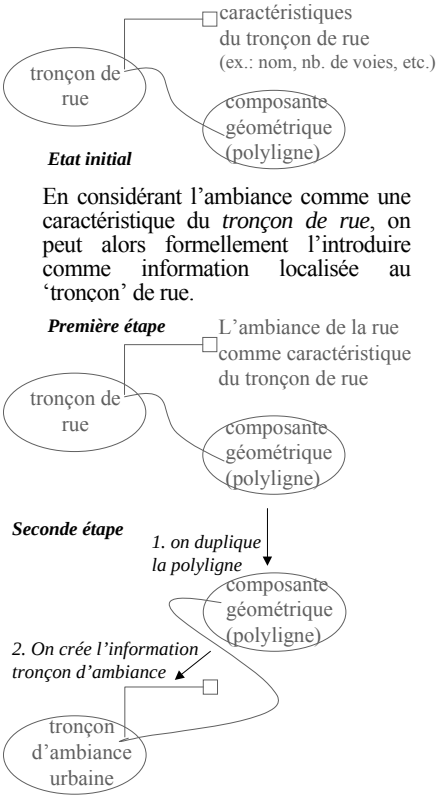
Figure C.21. *Extrait de la BD TOPO®*

En milieu dense, les bâtiments ne possèdent pas de contour individualisé. Leur composante géométrique est « amalgamée » à celles des bâtiments voisins.

C.4.2.2. Elaboration de l'objet méso-structurel

Pour appréhender l'ambiance urbaine nous avons cherché à définir une 'portion de territoire' pour agréger les 'objets micro' : 'commerces' (ex. : les commerces voisins, le commerces de même types dans une rue) et qui soit susceptible d'une plus grande homogénéité que la *rue.*

Dans un premier temps, nous avons considéré que l'ambiance pouvait être une caractéristique du *tronçon de rue*. Cette conception nous a permis de la localiser et de la transformer en information géographique : nous l'avons modélisée sous forme *d'information localisée* ayant pour support le 'tronçon de rue' (cf. p 106) (fig. C.22). Dans un second temps, nous avons identifié l'*emprise au sol* de l'ambiance à celle du *tronçon de rue* pour définir un 'tronçon d'ambiance' (fig. C.22 et C.23) .



En considérant l'ambiance comme une caractéristique du *tronçon de rue*, on peut alors formellement l'introduire comme information localisée au 'tronçon' de rue.

Figure C.22. *Elaboration de l'objet 'tronçon d'ambiance urbaine'*

Figure C.23. *Processus formel permettant de modéliser l'ambiance urbaine sous forme d'une information géographique*

Le 'tronçon d'ambiance' était un objet géographique proche de ce dont nous avions besoin, mais la primitive géométrique associée de type polyligne nous nous permettait pas de disposer d'une 'emprise' d'agrégation. Alors, dans une dernière étape nous avons considéré que l'ambiance était un phénomène diffus pour définir l''objet méso-structurel' comme un voisinage autour du 'tronçon d'ambiance urbaine'[19].

Suite à ce travail de formalisation technique, nous disposions pour mettre en œuvre la méthode d'observation multi-niveaux d'un 'objet méso' d'observation principal pour le suivi du phénomène avec un SIG (le voisinage du 'tronçon d'ambiance urbaine') et d''objets micro' pour construire des caractéristiques (les 'commerces'). Il ne restait plus qu'à formaliser la seconde étape de la méthode d'observation multi-niveaux : la *caractérisation*.

C.4.2.3. Quelles caractéristiques de l'ambiance observer ?

Pour créer des indicateurs de suivi, nous avons cherché comment caractériser l'ambiance, dans une perspective plus proche de celle du visiteur que du chaland. Le thème de *l'animation* permet d'en appréhender plusieurs aspects dans ce sens (fig. C.24). En effet, la notion d'animation (« *caractère de ce qui est animé, vif, plein de vie* » [Petit Robert, 2001]) est en relation avec le sentiment de *sécurité* (une rue animée n'est pas « morte », ni déserte ; elle paraît plus sûre) et avec la *convivialité* (le mot *animation* a pour synonyme : *chaleur*, *passion*, *entrain*). Elle fait également la liaison avec l'approche commerciale (grâce au lien entre *animation* et *activité-mouvement*). Pour caractériser 'l'ambiance urbaine d'un tronçon de rue' nous avons donc retenu le thème de l'*animation* que nous avons qualifié selon trois approches [Basile et Bordin, 2005] :

– l'*animation par l'activité commerciale*, dans une démarche proche des considérations de type géomarketing. Cette approche a révélé l'intérêt d'étudier en particulier les 'cellules vides'.

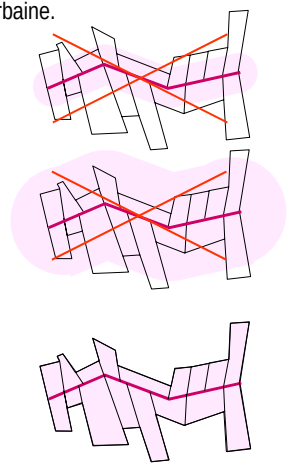
– l'*animation par le décor*, pour saisir des aspects plus qualitatifs de l'ambiance (comme la convivialité, la sécurité),

– l'*animation temporelle*, qui prend en compte les variations de l'*animation* (et donc de l'ambiance) dans le temps, mais à une échelle et selon un modèle temporel différant de ceux considérés pour le suivi et l'observation du phénomène.

L'animation par l'activité commerciale

L'animation par l'activité commerciale regroupe tout ce qui contribue à une animation thématique et notamment les types de commerces présents dans le 'tronçon de rue' étudié. Certaines activités marquent en effet en terme de standing et ou de spécialisation (une banque ou un bijoutier n'ont pas la même connotation qu'une centrale de téléphonie ou un sex-shop, de façon moins marquée des commerces d'alimentation n'ont pas le même « sens » que des commerces vestimentaires). Cette approche similaire à celles du géomarketing, utilise des références semblables, pour définir des indicateurs. En particulier, nous exploitons une typologie proche de celle employée par l'INSEE, entre autres, pour catégoriser les commerces, les codes NAF (Nomenclature des Activités Françaises). Elle permet de caractériser le 'tronçon de rue' en fonction de l'activité commerciale. Par exemple, elle permet de recenser le nombre de marchands de meubles (codes 36.1) et de préciser s'il y a des restaurants (code 55.3). Elle permet de regarder s'il existe une grande diversité de magasins ou au contraire si

19. En fait, la définition de l''objet méso-structurel à demandé un travail supplémentaire de modélisation. En effet, le voisinage a d'abord été envisagé sous forme de buffer (cf. p. 143), mais pour mieux correspondre à l'emprise réelle des boutiques dans les parcelles, nous avons décidé de construire un voisinage plus pertinent. Partant du raisonnement que les commerces qui font l'ambiance d'une rue sont ceux qui donnent sur la rue, nous avons alors défini un «voisinage topologique» : nous avons sélectionné toutes les parcelles qui avait une limite commune avec la 'rue' et nous les avons agrégées, pour définir le voisinage du tronçon d'ambiance urbaine.



Ce choix nous ramenait au cas b2, d'une information méso-structurelle sous forme d'une partition d''objets géographiques' en l'occurrence : les 'parcelles'.

Figure C.24. *L'animation comme caractéristique de l'ambiance d'un tronçon de rue*

elle est très spécialisée. Ainsi, nous avons constaté que l'une des deux zones géographiques choisies pour l'application (la rue Esquermoise, à Lille) possède une dominante 'marchands de meubles' (un peu à l'image du faubourg Saint-Antoire, à Paris). Les économistes ont montré l'avantage apporté dans certains secteurs par la spécialisation commerciale des tronçons de rue (en particulier pour les restaurants ou l'habillement {Sillard, 2003}).

Lorsque l'on considère l'animation par l'activité commerciale, les *cellules vides*, c'est à dire les boutiques sans activité (cf. figure C.25) sont des lieux sans offre pour le passant, et sans demande d'un point de vue économique, qu'il importe de relever. Elles deviennent alors le symptôme d'une situation critique (cadre urbain dégradé ou problèmes socio-économiques). Elles sont parfois décrites par des informations variées (ex. : durée de l'inactivité, activité avant fermeture, etc.). Ces informations sont alors autant d'éléments sur l'animation du 'tronçon de rue', sur son 'dynamisme économique'. Par leur apparence, ces cellules influent aussi sur les perceptions visuelles des passants et sur leur ressenti de la rue. Dans la continuité du parcours, elles apparaissent comme des « trous noirs », des «coins sombres». Elles nous ont ainsi servi également à appréhender l'animation par le décor.

Figure C.25. *Parmi les commerces, les' cellules vides' sont des objets symptomatiques pour l'ambiance.*

L'animation par le décor

L'animation par le décor regroupe ce qui contribue aux perceptions sensorielles et qui concourt à façonner une impression globale de l'ambiance plutôt qualitative (ex. : tranquille, agréable, conviviale, etc.). Dans ce registre, nous nous intéressons aux boutiques elles-mêmes (et non à leur activité, au contenant et plus au contenu), à leur mode d'insertion dans les bâtiments, à leur façade, etc. Notamment, nous qualifions l'agencement de leur vitrine au travers de leur système de fermeture plus ou moins opaque (fig. C.26). Les cellules vides réapparaissent ici comme le cas de fermeture de plus grande opacité.

Figure C.26. *L'animation par le décor au travers des rideaux de fermeture*

L'animation temporelle

Lorsque nous réfléchissons à l'observation temporelle de phénomènes géographiques pour proposer une méthode de suivi dans les SIG, nous envisageons généralement des échelles de temps longues[20] ; il s'agit de suivre les grands mouvements d'évolution, d'observer des transformations. Cependant, en appréhendant 'l'ambiance urbaine' au travers de l'*animation* d'un tronçon de rue, il est également pertinent d'introduire une observation temporelle sur des laps de temps plus courts, l'animation d'une rue variant parfois énormément d'un moment de la journée à l'autre (ex. : début de matinée/soir, midi/après-midi, semaine/week-end). Les variations temporelles deviennent alors une *caractéristique* de l'ambiance : une rue « animée jour et nuit » n'équivaut pas à une rue « déserte en dehors des heures de bureau ». Il s'agit ici d'observer ces variations d'animation pour tirer des caractéristiques sur le 'tronçon de rue'. Nous avons donc construit des indicateurs à partir des *heures d'ouverture des boutiques (cf. annexe 7)*. Le suivi des changements d'animation à l'échelle de la journée (ou la semaine) (fig. C.27), dans un modèle de temps cyclique (cf. paragraphe A.2.2.2.2), permet de caractériser l'animation et de qualifier ainsi le 'tronçon d'ambiance urbaine'. Par exemple, le tronçon de rue est animé le soir, ou le midi, il ne l'est pas le week-end, etc. Ce sont en revanche les variations de cette caractérisation dans un temps long et selon un modèle de temps linéaire qui servent au suivi du phénomène de *l'ambiance urbaine* (par exemple, le tronçon de rue s'anime de plus en plus le soir, de plus en

20. probablement influencé en cela par les délais de production et d'intégration des données géographiques.

plus le dimanche, de moins en moins tôt, etc.). Ce genre d'observation journalière (ou hebdomadaire) rappelle les études en *Time geography*. Des travaux dans ce domaine apporteraient probablement des éléments pour approfondir cette partie de notre étude. Par ailleurs, on pourrait introduire d'autres critères d'observation portant sur des indicateurs du même type, relatif à d'autres temps cycliques (par exemple le temps des saisons, [Pradel, 2005b]).

Figure C.27. *Variation de l'animation en fonction du temps dans un modèle temporel considéré comme cyclique*

C.4.3. Mise en œuvre de l'observation de l'ambiance urbaine

C.4.3.1. Propriétés d'un indicateur

Une étape importante pour observer des phénomènes géographiques dans un SIG consiste à construire des indicateurs de suivi. Pour que ces indicateurs soient exploitables, ils doivent posséder certaines propriétés (fig. C.28) :

– ils doivent être *transposables* dans le temps (pour observer différentes époques) et dans l'espace (pour considérer différents sites soumis au phénomène) ; l'existence d'un indicateur ne doit dépendre ni du lieu ni de la date.

– ils doivent être *reproductibles* d'une date à l'autre (pour observer des évolutions), d'un site à l'autre (pour comparer deux situations géographiques différentes) ; la forme d'un indicateur ne devrait pas dépendre du lieu ou de la date d'observation. Il ne suffit pas de proposer un indicateur construit à partir d'une typologie des commerces pour pouvoir comparer l'ambiance entre deux tronçons de rue, l'un en France, l'autre en Italie. Il faut également que les typologies employées ne soient pas spécifiques à chaque cas. Sans être forcément identiques, elles doivent être cohérentes, sinon toute comparaison (dans l'espace ou le temps) est impossible, limitant de fait toutes observations et analyses. Il s'agit ici de garantir la faisabilité de la méthode d'observation (en s'assurant de pouvoir suivre les variations de l'indicateur dans le temps) et de lui assurer un champ d'applications plus large (possibilité d'observer les variations de l'indicateur dans l'espace).

– enfin, dans le cas d'un phénomène subjectif, une contrainte supplémentaire s'impose : l'indicateur doit être robuste ; sa valeur ne doit pas dépendre de celui qui l'établit. Dans le cas contraire, il devient difficile de déterminer parmi les variations observées, celles qui correspondent à de véritables variations sur le terrain et celles dues à des différences interpersonnelles. Il s'agit ici de la qualité de la méthode d'observation.

propriétés de l'indicateur		**qualités de la méthode**
transposable	→	faisabilité
reproductible	→	comparabilité
robuste	→	Non biaisée

Figure C.28. *Propriétés des indicateurs*

C.4.3.2. Construction des indicateurs

Pour comparer des situations variées et contribuer à une meilleure compréhension du phénomène étudié, il importe de pouvoir étudier parallèlement l'*ambiance urbaine* sur des sites différents. Pour cette application nous avons d'ailleurs choisi deux zones d'application, avec pour objectif de tester l'apport de la méthode dans le cadre d'observations comparatives. Nous avons ainsi retenu deux rues : la rue Esquermoise à Lille (France) et la via della Maddalena à Gènes (Italie). Les deux sont situées en centre historique, mais reflètent des situations très différentes qui se perçoivent entre autre dans le discours d'acteurs privilégiés. Or les données utilisées proviennent en partie de sources (nationales) différentes ; Elles ne comportent pas les mêmes informations ou ne font pas toujours référence aux mêmes caractérisations. Un travail de préparation sur les

données a donc été nécessaire pour assurer la reproductibilité des indicateurs.

C.4.3.3. Données utilisées, éléments techniques

Certaines données ont été fournies par des organismes en charge de la gestion de l'activité commerciale (notamment par les chambres de commerce et d'industrie (CCI), mais aussi par des syndicats patronaux de commerçants). Ces données ont été complétées ensuite par des relevés directs sur le terrain. En particulier, concernant l'activité de chaque boutique, nous n'avions pas de données sur le site de Gènes et l'attribut utilisé dans les données sur Lille se référait à une codification ancienne et spécifique. Pour pouvoir transposer la méthode d'observation à d'autres sites et permettre des études comparatives, nous avons choisi de normaliser cette information. Une retranscription des informations existantes a ainsi été effectuée en nous appuyant sur le système européen de classement des activités économiques (qui se traduit en France par les codes NAF et en Italie par les codes ATECO). Cependant, pour éviter une logique d'observation uniquement économique, nous avons préalablement affiné la codification européenne pour intégrer des caractéristiques qui dépassent la simple typologie commerciale (par exemple, nous avons introduit la précision 'antiquaire' dans la rubrique 'marchand de meubles' en considérant que cette spécialisation apportait une valorisation supplémentaire).

Les enquêtes terrain sur les sites d'application ont également fourni les informations sur le *type de fermeture* et les *horaires d'ouverture*, qui, du fait de leur spécificité, n'existent dans aucune base de données. Pour les implémenter dans un SIG, nous avons travaillé sur leur modélisation en données géographiques ; en effet, au travers de cette modélisation, il s'agit d'avoir accès à une caractérisation plus *qualitative* de *l'ambiance urbaine* (ex. : esthétique des rideaux, type d'animation liée aux horaires d'ouverture), tout en assurant aux critères de caractérisation une objectivité suffisante pour garantir la *robustesse* de l'indicateur observé. Ce travail a conduit à proposer des éléments de spécification de saisie pour intégrer ces informations sous forme d'attribut des objets 'commerces':

– pour les horaires d'ouverture, plutôt que de choisir la forme objective consistant à donner l'information brute : heures d'ouverture/ fermeture, nous avons adopté une codification par tranches ou créneaux horaires en fonction des jours de la semaine. Cela permet d'identifier plus simplement les moments (de la journée et/ou de la semaine) où le tronçon de rue est le plus animé, car les occurrences principales sont plus facilement repérables. Cette modélisation permet d'introduire l'information sous forme plus applicative, avec une dimension plus interprétative, par exemple : avant 7h30 (très matinal), de 9-12h (matinée classique), après 19h (tardif), dimanche matin (loisir matinal). L'observation de *l'animation* permet d'analyser le fonctionnement de la rue : Est-ce un fonctionnement traditionnel (ex. : fermé le midi), ou bien est-il dans une mouvance plus moderne (ex. : service en continu) ? Est-ce un fonctionnement spécifique (ex . : ouverture le dimanche, horaire tardif) particulier au site ou bien le retrouve t'on ailleurs, ce qui pourrait alors correspondre à l'émergence de tendances nouvelles ? Cette information met en particulier en avant des éléments culturels : par exemple, en France, le lundi est un jour de fermeture pour les commerces, alors qu'en Italie (et à Gênes en particulier), les magasins dans le secteur alimentaire ferment généralement le mercredi.

– pour observer l'animation par le décor, les systèmes de fermeture nous paraissent des éléments pertinents. Ils peuvent servir à caractériser

Figure C.29. *Deux exemples de systèmes de fermeture plus ou moins conviviaux (ouvert avec éclairage intérieur ou fermé (code 2 et code 7)* http://www.blindalu.com/galva2.jpg

Figure C.30 *Fermeture par une grille qui laisse voir la vitrine par-devant (code 2)* http://www.technostor.fr/grimeta.htm

l'ambiance de la rue en horaire de fermeture (cf. fig. C.29 à C.31) en termes d'esthétique et de sentiment de sécurité. Pour entrer cette information dans un SIG, nous avons élaboré une typologie des fermetures. Nous les avons classées en fonction de leur capacité à laisser filtrer la lumière vers l'extérieur (pour le sentiment de sécurité) et à laisser ainsi entrevoir l'agencement de la vitrine (pour l'esthétique) : les rideaux les plus « transparents » ont été considérés globalement comme d'une plus grande convivialité. Nous avons appliqué cette typologie aux informations collectées sur le terrain.

Figure C.31. *Fermeture par une grille semi-fermée qui laisse voir partiellement la vitrine (code 3)*
http://www.ajcallendreau.com/serv4.htm

Sous réserve de données, les spécifications mises au point permettent d'observer avec un SIG des sites très différents (*observation de variations spatiales*). Pour observer des *variations temporelles* dans un SIG, l'utilisateur dépend beaucoup du mode d'intégration du temps dans les bases de données (cf. paragraphe A.3.2). Cependant, pour observer l'ambiance urbaine avec la méthode d'observation multi-niveaux, une modélisation simple comme celle des données utilisées est suffisante… à condition que les champs soient remplis de façon exhaustive. Or, ce n'était pas le cas. Les données étaient symptomatiquement lacunaires : les organismes qui recueillent et mettent à jour ces informations n'ont pas encore ressenti de façon suffisante le besoin d'inscrire dans le temps leurs observations. La collecte des données se fait parfois plus avec l'idée de péremption des données (les données sont-elles obsolètes?) et moins avec celle de suivi de l'observation. Pour ces organismes, il s'agit encore essentiellement d'une observation statique de type inventaire où l'enjeu est de disposer d'un état des lieux à jour. Il ne s'agit pas encore d'observer l'évolution de ces états pour dégager des tendances, dans une démarche plus prospective, dans une vision plus cinématique, voire dynamique pour étudier les causes de changements. Pourtant, l'analyse de la rotation des commerces (quel est le turn-over des commerces en fonction de leur localisation et de leur activité?) fournit des éléments de tendance intéressants qui pourraient être utiles pour la mise en place de politiques dédiées…
Ainsi, nous proposons aux collectivités une méthode pour mettre en place un outil de suivi de l'ambiance des centres villes ainsi que des éléments pour spécifier les bases de données associées. Cet outil peut servir à comparer des situations différentes. L'appropriation de cette application par les utilisateurs permettra de disposer du recul temporel pour effectuer des analyses temporelles plus pertinentes pour un même site ou entre sites différents.

C.4.4.Vers un outil opérationnel d'observation de l'ambiance urbaine au service des collectivités

Depuis une dizaine d'années, des opérations de « gestion de centre-ville » sont mises en œuvre. Elles visent à faire évoluer l'activité économique dans certains périmètres urbains limités, en transformant le cadre pour le rendre plus attractif pour les chalands. *L'ambiance* participe à ces opérations à la fois comme levier de transformation et comme résultant de ces opérations. L'objectif de cette application consiste à développer un outil pour observer le phénomène d'ambiance à l'aide d'indicateurs factuels, pour pouvoir disposer d'une base pour comparer les modèles des différents acteurs. Ces indicateurs envisagés dans une observation temporelle, doivent venir en complément des indicateurs commerciaux, pour évaluer ce que la nouvelle approche peut avoir apporté, pour voir si des changements d'ambiance et de perception se sont produits suite à l'introduction des nouveaux outils et préciser le cas échéant en quoi ils consistent. Par exemple, « comment les changements introduits se sont-ils greffés sur la ville préexistante ? ».

C.4.4.1. Enjeux d'un outil de suivi de l'ambiance urbaine pour la gestion de centre-ville

[Basile, 2005] présente l'enjeu applicatif d'un outil SIG opérationnel pour l'observation de l'ambiance urbaine. Il nous paraît double. Il est technique et stratégique :

– *technique* : il s'agit de transposer des méthodes issues du monde du marketing aux milieux opérant sur la ville. Or cette transposition ne peut pas être un simple transfert de compétences : le niveau d'expérience en SIG et en exploitation des données est très différent entre les deux « mondes »[21]. Par ailleurs, la particularité du phénomène géographique à observer demande un travail de modélisation conceptuelle spécifique : l'appréhension dans les SIG de *l'ambiance urbaine*, globale et intangible, telle que les collectivités aimeraient la saisir, représente en soi un défi. Il s'agit donc de concevoir une méthode qui permette aux utilisateurs d'observer le phénomène de l'*ambiance urbaine*, en leur évitant l'étape préliminaire de spécification, sans leur demander de renoncer à observer ce phénomène tel qu'ils le conçoivent au profit de modèles existants. Il faut que cette méthode soit peu coûteuse (en particulier qu'elle soit simple) à exploiter pour que les utilisateurs puissent se l'approprier, qu'elle soit compatible avec les méthodes du géomarketing complémentaires, enfin qu'elle soit « transportable » pour favoriser le partage d'expériences très important dans ce type d'application.

– *stratégique* : La démarche de gestion de centre-ville implique de nombreux acteurs. Il importe que tous puissent s'investir dans les aspects techniques, pour que les outils mobilisés répondent à leurs besoins et deviennent réellement pour chacun des supports à la réflexion. Les conditions d'un véritable dialogue passe par cette appropriation. Elle seule peut permettre d'éviter le risque d'une prévalence d'une logique professionnelle sur les autres (d'une logique commerciale mondiale en particulier). En effet, certains utilisateurs, contraints de déléguer l'exploitation des outils techniques à ceux qui les maîtrisent, prennent le risque que ces outils valorisent principalement les aspects intéressant leurs gestionnaires. Il s'agit donc de proposer une méthode qui soit simple à mettre en œuvre dans sa configuration initiale, qui permette d'élaborer des outils quasiment « clé en main » et qui soit susceptible d'évoluer pour s'adapter au niveau de compétence des utilisateurs.

Pour M. Basile, comme pour nous, cette application constituait avant tout, une question méthodologique. Tandis que pour nous il s'agissait de tester la méthode d'observation générique dans une mise en œuvre spécifique, pour Maria Basile, il s'agissait :

- *« de tester un nouveau regard sur l'ambiance »*, de valider son apport pour l'approche de gestion de centre-ville,
- de profiter de la capacité des SIG à croiser des informations de registres variés sur la base de leur localisation commune, pour « voir autrement », d'utiliser les SIG pour « poser » et croiser des données qu'il serait difficile de mobiliser simultanément autrement.
- *« Surtout la mise en place de cet outil observatoire constitue un support à la réflexion (il aide à « voir comment voir ») »*, inscrivant l'observation du phénomène à l'aide d'un SIG dans un processus d'études et d'analyses plus élaboré.

Or, si certaines réserves ont été émis concernant essentiellement la disponibilité des données (données de références, données descriptives et données temporelles) les résultats ont été jugés globalement concluant. Ce faisant l'intérêt de la méthode d'observation spécifique permet de valider sur cet exemple l'intérêt de la méthode générique.

21. Bien que faisant partie des applications « tardives » des SIG, le géomarketing qui s'est développé à partir de 1992 est devenu un domaine d'applications très actif.

C.4.4.2. Enjeux du suivi de l'ambiance urbaine pour les SIG

Cette application offrait un excellent cadre pour mettre en œuvre la méthode d'observation multi-niveaux. Outre qu'elle insiste sur le travail de conception nécessaire pour appréhender le phénomène étudié dans un SIG et qu'elle présente une configuration du problème nécessitant de tester deux formes particulières de la méthode (cf. paragraphe B.2. p.139), cette application illustre le potentiel des applications d'observation dans un SIG. En effet, dans cet exemple, le phénomène observé (*l'ambiance*), n'est pas le phénomène principal étudié (*la transformation des centres villes*). Dans ces travaux M. Basile s'intéresse en effet à l'impact de la gestion de centre-ville sur les transformations urbaines et *l'ambiance* est l'un des thèmes à partir duquel elle propose de l'aborder. Les indicateurs permettant de suivre l'un, servent alors à analyser l'autre, et par rétroaction à mieux prendre les décisions. Dans cette application, l'observation dans le SIG est loin d'être seulement une application de type inventaire, elle s'inscrit déjà dans un processus d'analyse et d'aide à la décision.

Dans ce cas d'étude par ailleurs, nous avons montré qu'il importe que la mise en œuvre de l'application soit simple. Cela justifie la contrainte imposée au départ, d'exploiter des fonctionnalités standard. Dans ce cas particulier, nous avons proposé une solution de mise en œuvre initiale qui exploite uniquement des fonctionnalités basiques du logiciel SIG Mapinfo (Mapinfo étant considéré plus comme un logiciel de bureautique que comme un logiciel technique).

Enfin, ce cas d'application illustre comment des recherches complémentaires, portant sur des préoccupations thématiques applicatives d'une part, techniques et conceptuelles d'autre part, peuvent se rejoindre et s'enrichir autour de la conception d'une méthode d'observation spécifique.

C.5. Conclusion sur les trois déclinaisons de la méthode d'observation générique

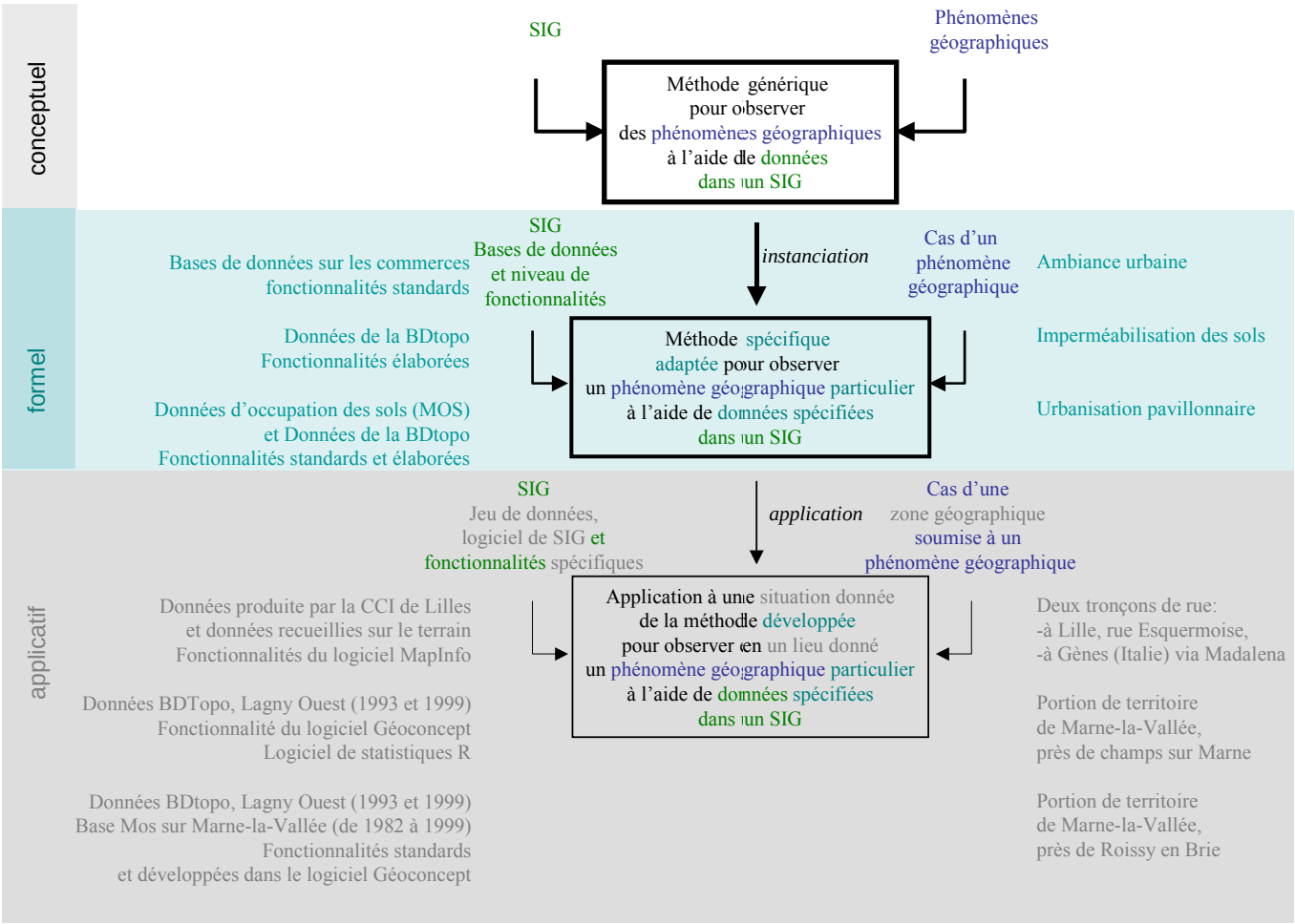

conceptuel
SIG
Phénomènes géographiques
Méthode générique pour observer des phénomènes géographiques à l'aide de données dans un SIG
formel
SIG
Bases de données et niveau de fonctionnalités
instanciation
Cas d'un phénomène géographique
Bases de données sur les commerces
fonctionnalités standards
Ambiance urbaine
Données de la BDtopo
Fonctionnalités élaborées
Imperméabilisation des sols
Données d'occupation des sols (MOS)
et Données de la BDtopo
Fonctionnalités standards et élaborées
Urbanisation pavillonnaire
Méthode spécifique adaptée pour observer un phénomène géographique particulier à l'aide de données spécifiées dans un SIG
applicatif
SIG
Jeu de données, logiciel de SIG et fonctionnalités spécifiques
application
Cas d'une zone géographique soumise à un phénomène géographique
Données produite par la CCI de Lilles
et données recueillies sur le terrain
Fonctionnalités du logiciel MapInfo
Deux tronçons de rue:
-à Lille, rue Esquermoise,
-à Gènes (Italie) via Madalena
Données BDTopo, Lagny Ouest (1993 et 1999)
Fonctionnalité du logiciel Géoconcept
Logiciel de statistiques R
Portion de territoire de Marne-la-Vallée, près de champs sur Marne
Données BDtopo, Lagny Ouest (1993 et 1999)
Base Mos sur Marne-la-Vallée (de 1982 à 1999)
Fonctionnalités standards
et développées dans le logiciel Géoconcept
Portion de territoire de Marne-la-Vallée, près de Roissy en Brie
Application à une situation donnée de la méthode développée pour observer en un lieu donné un phénomène géographique particulier à l'aide de données spécifiées dans un SIG

Figure C.32. *De la méthode générique à la mise en application pour l'observation de phénomènes spécifiques sur des zones particulières*

Cette partie de la thèse montre comment nous avons élaboré des méthodes d'observation pour suivre les trois cas de phénomènes que nous avions choisis d'étudier. Une application partielle a été mise en œuvre pour chacun des cas, pour tester auprès de thématiciens la pertinence des méthodes d'observation proposées (cf. figure C. 32).

[Basile et Bordin, 2005] et surtout le chapitre 2.2.2 de [Basile, 2005] (*« Un SIG pour accompagner la réflexion »*) présentent les résultats de la mise en application de la méthode spécifique pour l'observation de l'ambiance urbaine (dans l'Annexe 7, nous avons extrait quelques unes des cartes issues de cette application). La thèse de M. Basile conclue sur l'intérêt de cette mise en œuvre, même si elle relève que le manque de données reste aujourd'hui un handicap pour un suivi opérationnel.

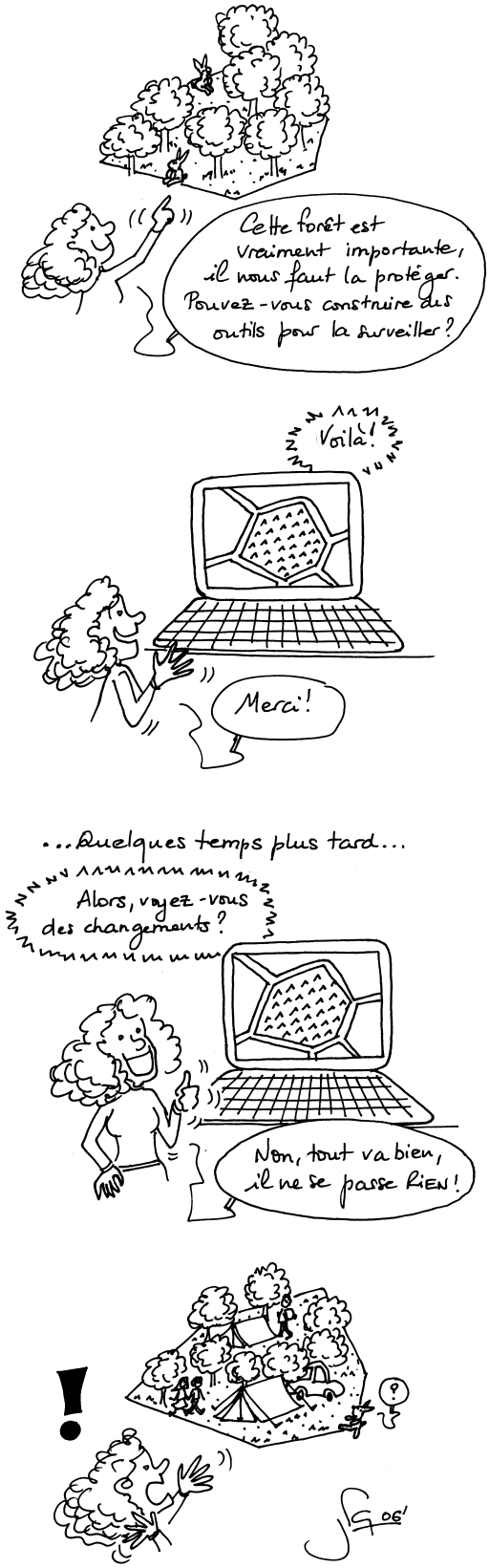
[Bordin et al., 2005] et surtout le rapport de stage[22] de [Mayeux, 2004] *« Suivi du coefficient d'imperméabilisation à l'aide d'un SIG »* co-encadré avec J-C. Deutsch, présente en détails la mise en application de la méthode de suivi du coefficient d'imperméabilisation (l'Annexe 7 présente un des

22. Stage scientifique de première année de l'Ecole Nationale des Ponts et Chaussées (ENPC) du 1/4/2004 au 1/7/2004

aspects non évoqués dans cette partie). Le rapport se conclue par : *« A l'issue du stage, nous pouvons affirmer que la méthode de calcul est prometteuse. Si les modèles doivent être améliorés et étendus à l'ensemble des éléments topographiques, le travail effectué montre néanmoins qu'un nouveau calcul, plus automatique, du coefficient d'imperméabilisation est possible».*

Le rapport de [Raimond, 2005] sur l'*« identification des zones d'habitat individuel à l'aide de données topographiques »* présente d'autres résultats d'application de la méthode d'observation générique. L'objectif du stage[23] était d'automatiser l'identification des 'zones d'habitat individuel' pour offrir aux utilisateurs un 'objet méso' géographique support pour suivre les phénomènes urbains qui s'y réfèrent. L'application développée sur une zone test de la BD TOPO® ont permis de saisir 4 types architecturaux classiques de 'zones d'habitat individuel' et de comparer deux méthodes pour les obtenir (voir en Annexe 7 deux extraits de cartes illustrant ces résultats). En annexe 7 nous présentons également d'autres éléments d'application au suivi des zones « pavillonnaires » qui nous ont servi à tester la méthode. Enfin l'annexe 8 propose une méthodologie pour mettre en œuvre les différentes étapes participant au développement d'une application de type suivi de phénomène à l'aide d'un SIG.

23. Stage de master 2 en Sciences de l'Information Géographique, Université de Marne-la-Vallée, du 1/3/2005 au 1/9/2005.


Cette forêt est vraiment importante, il nous faut la protéger. Pouvez-vous construire des outils pour la surveiller?
Voilà!
Merci!
...Quelques temps plus tard...
Alors, voyez-vous des changements?
Non, tout va bien, il ne se passe RIEN!
!
?

Dessin J-F Gleyze
Scénario P.Bordin 2006

Conclusion et perspectives

Conclusion

Le suivi de phénomènes dans un SIG

Le « suivi de phénomènes géographiques avec un SIG » soulève le problème de la gestion du temps dans ces outils. En effet, dans beaucoup de SIG, le temps n'est conçu que comme un jalon pour identifier les différentes versions des bases de données. Il est alors difficile de faire un suivi du territoire autre que visuel et comparatif ; automatiser l'observation des liens temporels entre les différents états d'une entité nécessite déjà souvent d'introduire des méthodes et des outils spécifiques. Tout comme il est complexe de travailler en 3D à partir d'objets mathématiques 2D (fig. D.1), il est délicat d'effectuer des *analyses dynamiques* dans un SIG, car elles requièrent de pouvoir faire le lien entre les causes et les effets, entre les événements (objets temporels) et les objets spatiaux et entre leurs caractéristiques respectives.

Figure D.1. *Comment concevoir la 3D dans un univers 2D ?* www.math.ttu.edu/current/tgtc2002/

Après avoir distingué plusieurs types de suivi de phénomènes (chapitre A.2), nous avons proposé de développer une méthode de suivi qui ne porte pas sur la réalisation d'analyses dynamiques – du moins pas directement – ; concevant le suivi comme une *« surveillance attentive et continue dans son cours »*, nous avons cherché une méthodologie pour élaborer des outils-SIG qui puissent servir à la fois de base de connaissances et de tableaux de bords, d'outils de comparaisons et de veille. Nous avons visé l'élaboration d'applications pour :

– produire des informations utiles à la construction, aux tests puis à la validation des modèles explicatifs,

– et ensuite pour mettre en évidence
- les spécificités locales (dans une perspective spatiale)
- l'apparition de faits nouveaux (dans une perspective temporelle).

Intégration des SIG dans les processus d'acquisition de connaissances sur les phénomènes géographiques

Ainsi dès l'origine de ces travaux de recherches, il nous a paru important de « faciliter l'intégration de la méthode dans les processus d'acquisition de connaissances sur les phénomènes géographiques étudiés ». Plusieurs raisons justifiaient à nos yeux cette importance, elles se sont traduites en objectifs spécifiques dans l'élaboration de la méthode de suivi :

– nous voulions faciliter l'intégration de la méthode dans le *processus de traitement* des informations ; il s'agissait déjà pour nous de fluidifier les échanges entre les différentes étapes participant aux traitements des informations (et entre les différents acteurs intervenant dans la mise en place d'un SIG d'observation). Il importe que les informations exploitées dans chaque étape soient au plus près de celles attendues par la suivante et de celles fournies par la précédente.

Nous avons donc commencé par identifier plusieurs étapes de formalisation. Parmi elles, nous avons distingué la *formalisation technique*, située entre la formalisation thématique et l'implémentation dans le SIG à l'aide de données, comme une étape stratégique. Après avoir présenté les conséquences de cette étape sur les applications, nous avons pu proposer des éléments de modélisation en information géographique favorables à une observation directe des phénomènes dans les SIG. En particulier nous avons montré l'importance de disposer d'une représentation explicite au bon niveau d'observation, avec une composante géométrique directe de type polygone ou polyligne. Le chapitre B1 qui présente ces points affirme les enjeux d'une collaboration entre thématicien et technicien dans la mise en place d'une solution SIG.

– nous souhaitions faciliter l'intégration dans le *processus d'acquisition de connaissances* et pour cela fournir des informations participant de cette acquisition.

Constatant dans le chapitre B.2, que les données disponibles dans la pratique sont souvent insuffisantes pour permettre un suivi détaillé des phénomènes étudiés, que les données sur les évolutions sont souvent pré-interprétées par les méthodes de suivi temporel dans les SIG avant même de parvenir aux thématiciens, nous avons développé une méthode pour produire des informations variées, simples à suivre dans le temps pour permettre à l'utilisateur de saisir des évolutions plus nombreuses et plus subtiles (chapitre B.3). Le choix de nous inspirer des méthodes d'observation multi-niveaux en géographie, en reprenant l'idée de créer des informations utiles à partir d'une observation à un niveau plus fin, nous a contraint à résoudre auparavant un autre problème technique classique des SIG : la gestion des représentations multiples. L'introduction d'une méthode d'agrégation originale adaptée à ces outils utilisant une emprise comme un emporte-pièce, nous a permis de dépasser cette difficulté pour proposer une méthode d'observation multi-niveaux dans les SIG. Cette méthode est de fait conceptuellement proche de celles mise en œuvre par les thématiciens. Dans le paragraphe B.2.5 nous ébauchons en particulier un lien, en montrant comment une application différenciée de cette méthode peut permettre d'introduire un niveau d'observation intermédiaire. Ce point pourra faire l'objet de développements ultérieurs.

– *Faciliter l'intégration* impliquait également pour nous de ne pas proposer de solution qui nécessite des investissements et des développements complémentaires importants préalablement à la mise en

œuvre. En effet, il s'agissait ici de faciliter l'intégration de façon très *pratique*. Les capacités des SIG permettent en effet aujourd'hui de proposer des solutions élaborées. Cependant leur technicité explique probablement en partie qu'elles ne soient pas aujourd'hui passées dans la pratique des thématiciens.

Après avoir passé en revue les solutions existantes en matière d'intégration de la dimension temporelle dans les SIG (chapitre A.3), nous nous sommes placés dans la situation la plus élémentaire, mais aussi la moins favorable : le modèle par *snapshots*. Dans ce modèle les versions des bases de données se présentent sous forme d'instantanés, sans offrir de liens entre elles, ni de possibilités d'extraction automatique des évolutions. Ce modèle correspond cependant à la situation de beaucoup d'utilisateurs, car il est à la fois simple à appréhender (car proche du modèle des éditions des cartes papier) et simple techniquement à mettre en œuvre dans un SIG. Nous avons alors visé à l'élaboration d'une solution méthodologique et conceptuelle qui exploite le potentiel des SIG tant en modélisation des informations géographiques (chapitre A .1), qu'en analyse à référence spatiale, et qui s'appuie principalement sur des fonctionnalités standards des SIG. La méthode d'observation proposée, dans son application élémentaire, ne nécessite de mettre en œuvre qu'une requête basique de type « est inclus dans ».

En effet, cette méthode d'observation généralise la pratique traditionnelle en cartographie thématique et les habitudes prises avec le suivi de phénomènes démographiques, où l'utilisateur utilise des entités géographiques stables dans le temps, pour supporter des informations descriptives diverses et évolutives. Il en est ainsi du traitement classique des données de recensements : les unités (souvent administratives (ex. : communes, départements, régions) ou les îlots, etc.) servent de support aux données sur la population, sur les entreprises, sur le logement, l'éducation, etc. Le fait d'utiliser toujours les mêmes unités de références permet de considérer les informations collectées comme des caractéristiques de ces unités, et par comparaison des valeurs consécutives d'un même attribut de pouvoir en observer les évolutions. Ce principe est celui exploité par l'équipe de {MacMaster,2006} pour développer le NHGIS (National Historical Geographic Information System). Résolvant de nombreux problèmes techniques (de généralisation entre autre) pour maintenir la cohérences du support et disposer d'une partition de référence commune pour tous les recensements effectués depuis 1790 aux Etats-Unis, ils proposent de faciliter aux utilisateurs les études d'évolutions de nombreuses caractéristiques démographiques sur une large période de temps [www.pop.umn.edu]. Cependant, un problème posé par l'utilisation de telles unités est lié au fait que les phénomènes étudiés ne se localisent pas forcément en fonction d'un découpage prédéterminé, induisant alors des distorsions dans les observations. C'est le cas par exemple de l'occupation des sols. Le MOS de l'IAURIF propose une solution originale pour observer ces évolutions sur la région parisienne, qui consiste à définir une partition adaptée et de la maintenir comme telle dans le temps (à l'aide de sous-partitions). Cependant, le modèle utilisé, s'il facilite les inventaires temporels et le suivi des évolutions des parcelles de référence, ne permet pas d'introduire des caractéristiques supplémentaires adaptées à l'analyse détaillée d'une occupation des sols spécifique.

La méthode générique d'observation multi-niveaux élaborée dans cette thèse propose d'aider les utilisateurs à mettre en œuvre des

méthodes d'observation qui comme pour le MOS conviennent au suivi des phénomènes particuliers qu'ils étudient et comme pour le NHGIS permettent d'intégrer des caractéristiques nombreuses et variées.

– Enfin, envisageant par ailleurs l'intégration des SIG dans le processus d'aide à la décision, nous voulions que les SIG soient capables ultérieurement de découvrir des informations 'nouvelles', qu'ils ne se contentent pas de restituer de façon élaborée les informations qui ont été confiées à sa gestion.

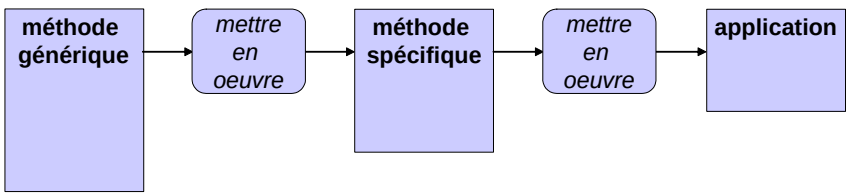
Nous avons ainsi dès le début de la thèse (chapitre A .1) situés les SIG par rapport aux *outils décisionnels*. Nous avons alors constaté que dans ces outils les fonctions de gestion des données, de restitution de ces données et de découverte de données nouvelles étaient assurées par des noyaux logiciels différents. Reprenant l'idée de cette modularité, nous avons conçu la méthode pour qu'elle reste *ouverte* sur les techniques d'analyses de données en datamining en tant qu'outils de « fouille de données » par excellence, mais également sur les méthodes et les outils d'analyses géométriques, la caractérisation de la morphologie et de la topologie constituant par ailleurs un apport singulier des SIG. L'analyse de ces deux composantes leur permet en effet de produire des caractéristiques spécifiques, pouvant être d'une très grande pertinence dans l'analyse de certains phénomènes comme le montre le cas particulier du phénomène d'urbanisation pavillonnaire… Puis, concevant cette ouverture

- comme le moyen de s'adresser à des utilisateurs de profils variés sans pour cela devoir compliquer l'« intégration pratique » de la méthode,
- mais également dans une perspective temporelle, comme une réponse possible à l'évolution des besoins des utilisateurs en terme d'analyse,

nous avons introduit en fin de chapitre B.2, consacré à la présentation de la méthode d'observation multi-niveaux, un paragraphe dédié aux extensions de la méthode par intégration de moyens supplémentaires et, en fin du chapitre B.3, un paragraphe dédié à l'intégration de modèles temporels plus sophistiqués pour saisir des informations d'évolutions plus pertinentes.

Perspectives

Ces travaux de recherches visent à aider ceux qui possèdent les compétences et les modèles thématiques pour étudier des phénomènes géographiques spécifiques ; ils proposent une méthode générale pour mettre en œuvre des outils SIG capables de fournir des données utiles à ces études. Les perspectives de cette thèse s'inscrivent pour nous dans la continuité de cet objectif.



Mise en œuvre de la méthode générique d'observation multi-niveaux pour l'observation d'un phénomène spécifique sur une zone particulière.

Pour les présenter nous proposons de reprendre le schéma de mise en œuvre de la méthode (figure ci-dessus).

Déclinaison de méthodes spécifiques et développement d'applications thématiques

Les premières perspectives sont celles ouvertes par la méthode elle-même : en collaboration avec des thématiciens, des déclinaisons de la méthode générique pourront être développées. A partir de ces méthodes spécifiques, certains utilisateurs pourront alors mettre en œuvre des applications pour étudier les phénomènes qui les intéressent sur des terrains particuliers.

Il pourra s'agir de cas nouveaux de phénomènes géographiques (ex. : suivi de l'urbanisation du littoral, développement de l'activité commerciale, étude de l'évolution du foncier), mais également de développer les trois cas présentés dans cette thèse. Il s'agirait de finaliser des applications. Dans ces recherches en effet, nous avons exploité ces cas principalement pour élaborer la méthode générique, pour nous confronter à la réalité, pour tester les hypothèses et la faisabilité des solutions envisagées, pour valider la méthode également, en la déclinant sur des cas particuliers, en fonction d'hypothèses sur le profil de l'utilisateur et de ces moyens. Nous avons ponctuellement mis en œuvre des éléments d'applications pour vérifier que les méthodes spécifiques permettaient bien d'aboutir à des résultats intéressants pour ceux qui étudiaient ces phénomènes. Cependant, pour disposer d'une présentation de la méthode plus aboutie, il faudrait intégrer ces résultats ponctuels dans des études thématiques « grandeur réelle » sur ces phénomènes.

Ces mises en oeuvre pourraient être l'occasion d'expériences nouvelles portant en particulier sur

- l'intégration de données variées, comme des données sur le parcellaire, des données INSEE (sur la population, sur les activités commerciales, etc.), des données de PLU (plan local d'urbanisme), etc.

- l'intégration de moyens d'analyses supplémentaires. Notamment, nous n'avons pas expérimenté l'apport des analyses topologiques. Or, elles nous paraissent particulièrement intéressantes, car elles établissent un lien entre les objets géographiques ; elles pourraient permettre de développer un nouveau type d'indicateurs, les *indicateurs de voisinage*, pour caractériser les objets en fonction de leurs voisins et en fonction de relations de proximité à d'autres objets (par exemple, un pavillon est séparé de plus de 300 m des ses voisins, une parcelle est entourée par deux parcelles appartenant au même propriétaire ou encore le pavillon est à moins de 1 km à pied d'une école et le trajet ne nécessite pas de traverser de rue).
Le développement de l'application sur l'ambiance urbaine pourrait également être l'occasion de faire un lien avec les méthodes et le savoir-faire développés pour le géomarketing.

- et dans une perspective « systémique », on pourra expérimenter l'intégration de ces expériences pour aider à l'observation d'autres phénomènes (le calcul du coefficient d'imperméabilisation des sols à partir du bâti pourraient en particulier, bénéficier d'une caractérisation morphologique fine des 'zones pavillonnaires').

Les éléments d'expériences tirés de ces mises en œuvre pourraient alors servir à améliorer la méthode elle-même.

Amélioration de la méthode et développements vers des analyses plus dynamiques

Nous avons insisté sur l'importance de proposer une méthode ouverte, en particulier ouverte sur les méthodes et outils d'analyses à références spatiales. Une piste d'améliorations consisterait à intégrer dans la méthode des éléments d'expériences extraits de ces autres domaines, en particulier, nous venons d'évoquer l'idée de développer des indicateurs sur les relations topologiques entre les objets. Ces caractéristiques supplémentaires viendraient s'ajouter à la gamme de celles pouvant être suivi avec la méthode d'observation. Mais l'intérêt principal, nous semble surtout dans le fait d'introduire une notion d'*observation du voisinage* qui pourrait servir à ébaucher des recherches pour des études plus dynamiques (Est-ce que les voisins ont changé ? Quels sont les changements de caractéristiques souvent consécutifs à un changement dans le voisinage ? Quelle est la probabilité que tel changement entraîne consécutivement un changement dans un voisinage (qu'il faudra préciser par ailleurs) et dans quels délais de temps ?, etc.).
Dans ce sens et en complémentarité, il faudrait étudier et formaliser l'intégration des modèles de gestion temporelle élaborés dont on a montré la capacité à fournir des informations intéressantes pour ce type d'analyses.

Par ailleurs, si l'on accepte l'hypothèse que la méthode d'observation sera capable de produire de nombreuses informations à observer, il serait intéressant d'établir des liens avec les outils de dataming, pour extraire des informations additionnelles et de suivre les recherches en « fouille de *données spatialisées* ».

En introduction nous citions [Hall, 1971] *: « Dans ce monde qui nous submerge d'informations, on comprend aisément pourquoi il est possible de se sentir perdre pied à l'intérieur de son propre domaine... D'où le besoin accru de structures de références susceptibles de faciliter l'intégration de la masse d'informations toujours changeantes que l'homme doit traiter ».*

En conclusion nous constatons que paradoxalement, notre objectif est de fournir plus d'informations encore,… mais de fournir également un outil pour les intégrer, pour les gérer, pour les traiter au mieux,… et même pour les suivre alors qu'elles sont toujours changeantes.
Nous pensons que les SIG pourraient y contribuer.

Bibliographie

[Abbott, 1998] Abbott E-A., *Flatland, a romance of many dimensions*, Penguin books, 1998, 144p.

[Allen; 1983] Allen JF., "Maintaining knowledge about temporal intervals", Communications of the *ACM*, 26, pp.832-843.

[Allen, 1984] Allen JF., "Toward a general theory of action and time", *Artificial Intelligence*, 23 (2), pp.123-154.

[Allen et Sanglier, 1979] Allen P-M., Sanglier M., "A Dynamic Model of Growth in Central Place Sytem", *Geographical Analysis*, 11, p. 256-272. 1979.

[Badard, 2000] Badard T., *Propagation des mises à jour dans les bases de données géographiques multi-représentations par analyse des changements géographiques*, Thèse de doctorat ès Sciences de l'Information Géographique, Université de Marne la Vallée, soutenue le 15 décembre 2000.

[Badard et Lemarié, 1999] Badard T, Lemarié C., "Prapagating updates between geographic databases with different scales" ; Chapter 10, *Innovations in GIS VII : GeoComputation*, Atkinson, P. and Martin, D. (Eds), Taylor and Francis, London, 1999, 12P ;

[Banos et Thevenin, 2003], Banos A., Thevenin T., « Révéler les rythmes urbains quotidiens », *6ièmes rencontres de ThéoQuant*, Besançon, Février 2003.

[Bararino, 2005] Bararino N., *De la qualité de vie au diagnostic urbain : vers une nouvelle méthode d'évaluation. Le cas de la ville de Lyon*, Université Lumière Lyon 2, Thèse en géographie, 8/6/2005.

[Bard, 2004] Bard S., *Méthode d'évaluation de la qualité de données géographiques généralisées, application aux données urbaines*, Thèse de doctorat de l'université Paris 6, soutenue le 26/1/2004.

[Barros, 2003] Barros J., Simulating Urban Dynamics in Latin American Cities, Proceedings of the *7th International Conference on GeoComputation*, Southhampton, 2003.

[Basile et Bordin, 2005] Basile M., Bordin P., « Méthode d'analyse de l'ambiance urbaine à l'aide d'un SIG », *Colloque REIT*, Liège, Belgique 2005.

[Basile, 2005] Basile M., *Urbanisme, management et commerce, la transposition des modèles belge et britannique de gestion de centre-ville dans les quartiers historiques des agglomérations de Gênes et de Lille,* Thèse de doctorat ès Aménagement et urbanisme, Université de Paris 8, 17 décembre 2005.

[Batty, 2001] Batty M., "Polynucleated Urban Landscapes", *Urban Studies*, vol. 38, 4, p635-655, 2001

[Bédard, 1999] Bédard Y., "Visual modelling of spatial databases: towards spatial PVL and UML", *Géomatica*, vol. 53, N°2, 1999, pp. 169-186.

[Bel Hadj Ali, 1997] Bel Hadl Ali A., *Appariement géométrique des objets géographiques et études des indicateurs de qualité*. Mémoire de DEA Sciences de l'information géographique, Université de Marne-la-Vallée, 1997, 109p.

[Benenson et Torrens, 2004] Benenson I., Torrens P., *Geosimulation : Automata-based Modeling of Urban Phenomena*, Wiley, 2004.

[Berry, 1964] Berry B.J.L., "Cities as systems within systems of cities", Papers of the Regional Science Association, 1964.

[Boffet, 2001 Boffet A., *Méthode de création d'informations multi-niveaux, pour la généralisation cartographique de l'urbain*, Thèse de doctorat en Sciences de l'Information Géographique, soutenue le 13/12/2001.

[Bonin,2002] Bonin O., *Modèle d'erreur dans une base de données géographiques et grandes déviations pour des sommes pondérées; application à l'estimation d'erreurs sur un temps de parcours*, Thèse de doctorat ès Mathématiques, Université Paris 6, 1/3/2002.

[Bonnefoy et Goiffon, 2003], Bonnefoy J-L., Goiffon M., Un système multi-agents pour modéliser le « retournement de l'espace » martiniquais, *6ièmes rencontres de ThéoQuant*, Février 2003.

[Bordin,2006] Bordin P., « De la carte à l'ordinateur pour observer le territoire », numéro spatial de *la Recherche*, n°398, juin 2006, p.38.

[Bordin, 2005a] Bordin P., « SIG et analyse de phénomènes évolutifs : Méthode de suivi de phénomènes évolutifs à l'aide de données élémentaires », *International Conference on spatial analysis and geomatics, SAGEO'2005,* Avignon, France, 20-23/6/2005.

[Bordin, 2005b] Bordin P., « Différenciation des informations géographiques pour améliorer l'analyse spatiale dans un SIG », 7^{ième} rencontre de *ThéoQuant 2005*, Besançon, 24-26/1/2005.

[Bordin et al, 2005] Bordin P., Mayeux M., Bonin O., Deutsch J-C., "Linking imperviousness coefficient to urbanism feature through GIS", *10th International Conference on Urban Drainage*, Copenhagen, Denmark, 21-26/8/2005.

[Bordin, 2002] Bordin P., *SIG, concepts, outils et données*, Hermès-Lavoisier, Paris, 2002, 256p.

[Bouillé, 1977] Bouillé F., Structuring cartographic data and spatial processes with the Hypergraph-Based Data Structure, *Advanced Study Symposium on Topological Data Structures in Geographic Information Systems*, MIT, October 17-21, 22p. in Proceedings,Vol.1.

[Braun, 2004] Braun A., *From the schema matching to the integration of updating information into user geographic databases*, Proceedings of the 12th International Conference on Geomatics, University of Gävle, Sweden, 7-9 june 2004.

[Brunet et al, 1992] Brunet R., Ferras R., Théry H., *Les mots de la géographie, dictionnaire*, collection dynamiques du territoire, Reclus critique, La documentation française, 1992.

[Bucaille, 1996] Bucaille O., *Mise à jour d'une base de données géographiques- Exemple de la BDCarto®.* Mémoire d'ingénieurCNAM, France, 1996, 119p.

[Bullock et al. 1974] Bullock N., Dickens P., Shapcott M., Steadman P., Time budgets and models of urban activity patterns, Social Trends, 5, pp.45-63.

[Bura et al, 1996] Bura S., Guerin-Pace F., Mathian H., Pumain D., Sanders L., "Multi-agents system and the dynamics of a settlement system", *Geographical Analysis*, vol.28, n°2, p.161-178, 1996.

[Cheylan et al, 1994] Cheylan J-P., Lardon S., Mathian H., Sanders L., Les problématiques liées au temps dans les SIRS, *Revue Internationale de Géomatique*, Vol. 4, No.3-4, 1994, pp 287-305.

[Chocat, 1997], Chocat B., coordonnateur de l'*Encyclopédie de l'hydrologie urbaine et de l'assainissement*, Lavoisier, TEC et DOC, p.1123, 1997.

[Chrismann, 1998] Chrismann N., Beyond the snapshot: Changing the Approach to Change, Error, and process, in (Eds) Egenhofer M-J., Golledge R-G., *Spatial and temporal reasoning in geographic information systems*, Oxford University Press, 1998, pp. 85-92.

[Chrismann, 1975] Chrismann N.R., Topological information systems for geographic representation, Proceedings of 2nd International Symposisum on Computer-Assisted Cartography, pp. 346-351.

[Claramunt, 2005] Claramunt C., « Syntaxe spatiale : vers de nouvelles formes de représentation d'une certaine logique de la cité », *ThéoQuant 2005*, Besançon, 24-26/1/2005.

[Claramunt et al, 1994] Claramunt C., De Sède M.H, Prélaz-Droux R, Vidale L., « Sémantique et logique spatio-temporelle, Revue internationale de Géomatique, vol. 4, n°2, 1994, pp.165-180.

[CODATSI, 1996] CODATSI, *Des données localisées aux systèmes d'information géographique, conseils aux services*, MELT, CCIG, 1996.

[Corcuff, 2005] Corcuff M-P., Automates cellulaires dynamiques et simulations de phénomènes directionnels, *7ième rencontre de ThéoQuant*, Besançon, 2005.

[Cornaut, 2000] Cornaut M., Reconquête et évolution de l'espace pavillonnaire, in Mémoire et projet du pavillonnaire en Nord-Essonne, *Cahier de la maison de banlieue N°3*, Centre culturel d'Athis-Mons, Mars-Avril 2000.

[Couclelis, 1999] Couclelis H., Space, time, geography, *Geographical Information Systems: principles, techniques, management, and applications*, 2nd ed., (edsLongley P., Goodchild M., Maguire D., Rhind D., Vol.1, New-York, Wiley, 1999, pp.29-38.

[Couclelis, 1998], Couclelis H., "Aristolian Spatial Dynamics in the age of geographic information systems", Chapitre 8, Egenhofer M.J., et Golledge R.G., *Spatial and temporal reasoning in geographic information systems*, Oxford, 1998, pp.109-118.

[Couclelis, 1994] Couclelis H., Toward an operational typology of geographic entites wil ill-defined boundaries, Position paper at the *workshop on objects with indeterminate boundaries*, TU Wien, 1994.

[Couclelis, 1992] Couclelis H., People manipulate Objects (but cultivate fields): beyond the raster-vector debate in GIS, in Frank, A.U., Campari I., Formentini U., (Eds), Theories and Methods of Spatial-Temporal Reasoning in Geographic Space, Berlin, Springer-Verlag, pp. 655-77.

[Cova et al., 2000] Cova T.J., Church R.L., Goodchild M.F., Extending geographic representation to include fields of spatial objects. First International Conference on Geographic Information Science (GIScience 2000), Savannah, Georgia, October,2000.

[Cuenin, 1972] Cuenin L., *Notions générales et principes d'élaborations, Cartographie générale*, Tome 1, Eyrolles, Collection scientifique de l'IGN, 1972.

[Dargemond 1984] Dargemond J., A classification of soft ware components commonly used in geographic information systems, in *Proceedings of the US/Australia Workshop on Design and implementation of Computer-Based Geographic Information Systems*, 1984.

[David et Fasquel, 1997] David B., Fasquel P., « Qualité d'une base de données géographique : concepts et terminologie », *BI de l'IGN*, n°67, 1997.

[Davis et Wallbridge, 1992] Davis M., Wallbridge D., « Winnicott, introduction à son œuvre », PUF, 2005, 192p.

[De Blomac, 2004] De Blomac F., « Faut-il mettre les SIG en observation ? », *SIG la lettre*, décembre 2004.

[Decrop, 2001] Decrop J., « Modélisation spatio-temporelle de la progression des surfaces résidentielles », *Vième rencontre de ThéoQuant*, Février 2001, Besançon.

[Dell'Erba et Libourel, 1997] Dell'Erba E., Libourel T., « Evolution d'entités géoréférencées ». Actes des *13ièmes journées Bases de Données Avancées (BDA'97)*, Grenoble, France, Septembre 1997, pp.89-110.

[Denègre et Salgé, 1996] Denègre J., Salgé F., *Les systèmes d'information géographique*, PUF, collection Que sais-je?, n°3122, 1996.

[Denis, 1998] Denis E., « Croissance urbaine et dynamique socio-spatiale ». Le Caire de 1950 à 1990, *L'espace géographique*, 2-1998, p129-142.

[Deutsch, 1989] Deutsch J-C., *Mémento sur l'évacuation des eaux pluviales*, La documentation française, Paris, 1989, 349p.

[Devogele et al, 1998] Devogele T., Parent C., Spaccapietra S., "On spatial database integration". *International Journal of Geographical Information Science* (IJGIS), 12(4), P. Fisher, M. Amstrong et B. Lees (EDS), June 1998, Taylor and Francis, London, pp. 335-352.

[Devogele, 1997] Devogele T., Processus d'intégration et d'appariement de bases de données géographiques – Application à une base de données routières multi-échelles. Thèse de doctorat ès Informatique, Université de Versailles, 12/12/1997, 207p.

[Didier et Bouveron, 1993] Didier M., Bouveron C., *Guide économique et méthodologique des SIG*, Hermès, 1993.

[Didier, 1990] Didier M., *Utilité et valeur de l'information géographique*, Economica, 1990.

[Duchêne, 2004], Duchêne C., *Généralisation Cartographique par Agents Communicants : le modèle CartACom*, Thèse de doctorat ès Informatique, Université Paris 5, 2004

[Duchêne et al, 2003] Duchêne C., Bard S., Barillot X., Ruas A., Trévisan J., Holzapfel F., « Quantitative and qualitative description of building orientation », 5th Workshop of ACI Commission, a generalization progress in automated map generalization, Paris, April 2003.

[Durand-Dastès, 2005] Durand-Dastès F., Note de synthèse sur la notion de phénomène, (inédit), Paris, 2005.

[Durand-Dastès, 1999] Durand-Dastès F., «L'explication en géographie » (1974, inédit), dans *Systèmes et modèles*, CD Rom , UMR Géographie-cités, Paris, 1999.

[Donnadieu et al, 2003] Donnadieu G., Durand D., Neel D., Nunez E, Saint-Paul L., *L'approche systèmique : de quoi s'agit-il ?*, Synthèse des travaux du Groupe AFSCET, « diffusion de la pensée systémique » septembre 2003.

[Egenhofer et Golledge, 1998] Egenhofer M-J., Golledge R-G., (Eds) *Spatial and temporal reasoning in geographic information systems*, Oxford University Press, 1998, 276p.

[Egenhofer et Herring, 1990] Egenhofer M.J., Herring J.R., A mathematical framework for the definitino of topological relationships. In Proceedings of the 4th International Symposium on Spatial Data Handling (SDH'90), Vol.2, Zurich, Switzerland, 1990, pp. 803-813.

[Ellerkamp 2001] Ellerkamp P., « Simulation des dynamiques spatiales de densification de la population autour d'Avignon », *Vième rencontre de ThéoQuant*, Février 2001, Besançon.

[Engelen, 2005] Engelen G., "Spatial simulations with cellular automata, recent advances in geography", *7ième rencontre de TheoQuant* , Besançon, 24-26/1/2005 .

[Ernst, 1978] Ernst B., *Le miroir magique de M.C. Escher*, Taschen, Köln, 1978.

[Essevaz-Roulet, 2005], Essevaz-Roulet M., « La mise en œuvre d'un SIG dans les collectivités territoriales », Dossier expert, *Urbanisme*, n0283, mars 2005.

[Feyt, 2004] Feyt G., Les métiers du territoire face aux technologies de l'information géographique: Babel et esperanto, Chapitre 2, in *Aspects organisationnels des SIG*, ss Dir. S. Roche et C. Caron, Hermès, 2004, 298p.

[Frank, 1998], Frank A.U., Different types of "time" in GIS, Chapitre 3, Egenhofer M.J., et Golledge R.G., *Spatial and temporal reasoning in geographic information systems*, Oxford, 1998, pp.40-62.

[Fritsch, 1997] Fritsch E., *Représentation de la géométrie et des contraintes cartographiques pour la généralisation du linéaire routier*, Thèse de doctorat, Université de Marne-la-Vallée, 1997.

[Georget, 2000] Georget E., Annuaire 1999 des SIG du METL, Exploitation annuelle du fichier de signalement, Collection de CERTU, Edition publique, 2000.

[Girard, 2004] Girard E., *Usage de la cognition spatiale pour localiser les lieux d'activité lors d'une enquête Origine-Destination*, mémoire de recherche à l'université de Laval, sous la direction de M. Theriault www.these.ulaval.ca/2004//21867/21867.html, (page consultée le 20/9/2006).

[Gleyze, 2005] Gleyze J-F., La vulnérabilité structurelle des réseaux de transport dans un contexte de risque, Thèse de doctorat en Analyse théorique et épistémologique en géographie, Université Paris VII, 2005, 848p.

[Guerin, 1968] Guerin J., *L'évolution de quelques tissus pavillonnaires de l'Est parisien*, Centre de recherche de l'urbanisme de Paris, 1968.

[Haggett et al., 1977] Haggett P., Cliff A.D., Frey A., Locational Models, Halstead Press, New-York, 1977.

[Hall, 1971] Hall EDT., *La dimension cachée*, Editions du Seuil, Paris, 1971, 255p.

[Hernandez, 1995] Hernandez D., HCI Aspects of a framework for the Qualitative Representation of Space, in *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*, (Eds) T.L. Nyergues, D.M. Mark, R. Laurini, M.J. Egenhofer, NATO ASI Series, 1995, pp.45-60.

[Holm et al, 2004] Holm E., Holme K., Mäkila K., Mattson-Kauppi M., Mörtvik G., "the microsimulation model SVERIGE; content, validation and applications', *SMC*, Kiruna, Sweden, 2004.

[Jamet, 1998] Jamet O., *Comparaison Vecteurs-Images pour la détection de changements des bâtiments d'une base de données topographiques*, Thèse de doctorat ès Signal et Images, ENST, Paris France, 11/12/1998, 204p.

[Jung, 2002] Jung F., Détection de changement par comparaison de couples stéréoscopiques, B.I N°73, *Bulletin d'Information de l'IGN*, La recherche en 2001 IGN (Institut Géographique National) p. 27-35, 2002

[Kelmelis, 1998] Kelmelis J.A., Process Dynamics, Temporal extend, and Casual Propagation as the basis for Linking Space and Time, in Egenhofer M-J., Golledge R-G., (Eds) *Spatial and temporal reasoning in geographic information systems*, Oxford University Press, 1998, pp. 94-104.

[Lalande, 1997] Lalande A., *Vocabulaire technique et critique de la philosophie*, Edition PUF, 1997 (première édition 1926) Articles : phénomènes, phénoménalisme, phénoménologie)

[Langran, 1993] Langran G., *Time in geographic information systems*, Taylor et Francis, London, 1993, 190p.

[Langran et Chrisman, 1988] Langran G., Chrismann N., "A framework for temporal geographic information" . *Cartographica*, 25(3), 1988, pp. 1-14.

[Latour et Le Floc'h, 2001] Latour P., Le Floc'h J., *Géomarketing*, Editions d'organisation, Mai 2001, 271p.

[Lemarié et Raynal, 1996] Lemarié C., Raynal L., Geographic data matching : First investigations for a generic tool. In Proceedings of GIS/LIS'96, Annual Conference and Exposition, Denver, Colorado, USA, November 19-21, 1996, pp. 405-420.

[Li *et al*, 1999] Li J., Jing N., Sun M., "An object-oriented data model integrating field and object, Geoinformatics and Socioinformatics", *the Proceedings of geoinformatics'99 Conference*, Ann Arbor, 19-21 June 1999, p1-10.

[Madoré, 2001], Madoré F., "New growth and territories", *The third congress on proximity*, 13-14/12 2001.

[Maguire *et al*, 1991] Maguire D.J., Goodchild F.M., Rhind D.W., *Geographical information systems, Vol.1: Principles*, Longman Scientific and Technical, Essex, England, 1991.

[Mathian et Piron, 2001] Mathian H., Piron M., Echelles géographiques et méthodes statistiques multidimensionnelles, Chapitre 2 in *Modèles en analyse spatiale*, Dir. L. Sanders, Hermès-Lavoisier, 2001, 330p.

[Mangin, 2006] Mangin L., L'avenir dans le dos, in *Pour laScience*, n°347, septembre 2006, 19p.

[Mayeux, 2004] Mayeux M., *Suivi du coefficient d'imperméabilisation à l'aide d'un SIG*, Stage scientifique de première année ENPC, Laboratoire COGIT, 2004.

[Millot, 2003] Millot M., *Développement urbain et insécurité routière : l'influence complexe des formes urbaines*, Thèse de doctorat, Ecole Nationale des Ponts et Chaussées, 2003

[Moeckel et al., 2002] Moeckel R., Schürmann C., Wegener M.,"Microsimulation of urban land use", Proceedings of *the 42nd Congress of the European Regional Science Association* (ERSA), Dortmund, 2002.

[Mustière, 2001] Mustière S., *Apprentissage supervisé pour la généralisation cartographique*, Thèse de doctorat en infromatique, de l'Université de Paris 6, soutenue le 8/6/ 2001

[Nicolescu, 1996] Nicolescu B., *La transdisciplinarité- Manifeste*, Editions du rocher, col. « transdisciplinarité », Monaco 1996.

[Nyerges et al, 1995] Nyergues T.L., Mark D.M., Laurini R., Egenhofer M.J., *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*, (Eds) NATO ASI Series, 1995, 435p.

[Nyman, 1998] Nyman L.W., "What Is GIS Anyway?", GISWorld, vol. 11, n°2, P.78, Janvier 1998.

[Ott et Swiaczny, 2001] Ott T., Swiaczny F., *Time-integrative geographic information systems, management and analysis of spatio-temporal data*, Springer, 2001.

[Ourednik, 2005] Ourednik A., Cellular segrgation toward a retroactive socio-spatial segregation model, *7ième rencontre de ThéoQuant 2005*, Besançon, 24-26/1/2005.

[Pastor, 2004] Pastor J., Conception d'une légende interactive et forable pour le SOLAP, Université Laval, faculté de foresteris et de géomatique, Maîtrise en sciences géomatiques, 2004.

[Petit, 2000], Mémoire et projet du pavillonnaire, in Mémoire et projet du pavillonnaire en Nord-Essonne, *Cahier de la maison de banlieue N°3*, Centre culturel d'Athis-Mons, Mars-Avril 2000.

[PetitRobert, 2001] Le Petit Robert, *dictionnaire de la langue française*, éditions Le Robert/Seuil, 2001.

[Peuquet, 2002] Peuquet D., *Representations of Space and Time*. Guilford, New York, 2002.

[Peuquet *et al*, 1998] Peuquet D., Smith B., Brogaard B., The ontology of Fields, report of specialist meeting held under the auspices of the varenius Project. *Talks and discussions at the specialist meeting*. Bar Harbor, Maine, June 11-13 1998

[Peuquet, 1986] Peuquet D., The use of spatial relations to aid spatial database retrieval, *Proceedings of the 2nd International symposium on Spatial Data Handling*, pp. 459-471.

[Pornon, 1990] Pornon H., *Systèmes d'information géographique, des concepts aux réalisations*. Hermès, 1990.

[Poupeau et Bonin, 2006] Poupeau B., Bonin O., 3D analysis with high-level primitives: a cristallographic approach, *Proceedings of SDH'06*, Vienna, 2006.

[Pardel, 2005a] Pradel B., Entre institutionnalisation et clandestinité : le graffiti ou l'hydre à deux têtes in "C'est ma ville!", l'Harmattan, 2005

[Pradel, 2005b] Pradel B., Polarité d'animation cyclique de l'espace public, Mémoire de DEA Mutations urbaines et gouvernance territoriale, Université de Marne-la-Vallée, 2005 ;

[Proulx et al, 2002] Proulx M-J., Suzie Larrivée, Yvan Bédard, Représentation multiple et généralisation avec UML et l'outil Perceptory, in *Généralisation et représentation multiple*, Dir. A. Ruas, 2002, Hermès, Lavoisier.

[Pumain, 2004] Pumain D., Analyse spatiale, généralités, *Hypergéo*, GDR Libergéo-http://hypergeo.free.fr

[Pumain et Saint-Julien, 1997] Pumain D., Saint-Julin T., *L'analyse spatiale, localisations dans l'espace*, Armand Collin, Paris, 1997.

[Quodverte, 1994] *Cartographie numérique et information géographique*, Thèse de doctorat, Université d'Orléans, 1994.

[Raimond, 2005] Raimond O., *Identification des zones d'habitat individuel à l'aide de données topographiques*, Mémoire de master 2 en Sciences de l'Information Géographique, Université de Marne-La Vallée-ENSG, Septembre 2005.

[Raymond et al, 2002] Raymond H., Haumont N., Raymond M.G., Haumont A., *L'habitat pavillonnaire*, L'harmattan, Paris, 2002.

[Raynal, 1997] Raynal L., « Gestion du temps dans une base de données topographiques », *bulletin d'information de l'IGN*, n°68, p.65-74, 1997

[Regnault et al, 1999] Regnault N., Edwards A., Barrault M., Strategies in building generalisation, *19th International Workshop on Automated Map Generalization*, Ottawa, Canada, 1999.

[Regnault, 1998] Regnault N., *Généralisation du bâti : Structure spatiale de type graphe et représentation cartographique*, Thèse de doctorat ès Informatique, Université de Provence, 1998.

[Reuchlin, 1990] Reuchlin M., *Psychologie*, PUF, huitième édition, 1977,679p.

[Rey, 1998] *Les territoires centre-européens, Dilemmes et défis*, L'Europe médiane en question, Lectio géographie, (Eds) V. Rey, 1998, pp. 14-19.

[Rey, 1996] Rey V., Un espace entre deux, in Europes orientales, Géographie, (Eds) R. Brunet, Belin-Reclus, 1996, pp.10-28.

[Ribardière, 2005] Ribardière A., *Précarité sociale quand les mailles s'en mêlent* , Thèse de doctorat ès Géographie, Université de Paris I, 15/12/2005

[Ruas et Mustière, 2005] Ruas A., Mustière S., Bases de données géographiques et cartographiques à différents niveaux de détails, Le monde des cartes – *Revue du Comité Français de Cartographie*, n°185, 2005, pp. 5-14.

[Ruas et Bianchin, 2002] Ruas A., Bianchin A., Echelle et niveau de détail, in A. Ruas (Ed) *généralisation, International Journal on GIS*, vol.12, n08, p. 789-803.

[Ruas, 2002] Ruas A., *Généralisation et représentation multiple*, Information géographique et aménagement du territoire, Hermès- Lavoisier, 2002.

[Ruas, 1999] Ruas A., *Modèle de généralisation de données géographiques à base de contraintes et d'autonomie*, Thèse de doctorat ès Sciences de l'Information Géographique, Université de Marne-la-Vallée, 1999.

[Sanders, 2006] Sanders L., Les modèles agent en géographie urbaine, in Amblard F., Phan D. (dir),. *Modélisation et simulation multi-agents pour les Sciences de l'Homme et de la Société : une introduction*, Hermès-Lavoisier, 2006, pp. 151-168.

[Sanders, 2001] Sanders L. (Eddir.), 2001, *Modèles en analyse spatiale*, Hermès-Lavoisier, 2001, 333p.

[Saint-Joan, 1996] Saint-Joan D., Raisonnement incertain et flou pour l'aide à la décision dans les systèmes d'information géographique, *Inforsid'96*, Bordeaux, 4-7 juin 1996.

[Saint-Julien, 2001] Saint-Julien T., Processus de diffusion spatiale et modélisations du changement, Chapitre 5 in *Modèles en analyse spatiale*, Dir. L. Sanders, Hermès-Lavoisier, 2001, 330p.

[Serres et Farouki, 1998] Serres M., Farouki N., *Le trésor, Dictionnaire des sciences*, Flammarion, 1998

[Schelling, 1971] Schelling T.S., "Dynamic Model of Segregation", Journal of Mathematical Sociology, 1(2), 1971, pp.143-186.

[Sheren et al, 2004] Sheren D., Mustière S., Zucker J-D., How to integrate Heterogeneous Spatial Databases in a Consistent Way?, In *Conference on Advanced Databases and Information Systems* (ADBIS), Budapest, 2004, paru en tant que LNCS3255, Benazue A., Demetrovics J., et Gttlob G (Eds), pp.235-245.

[Shiyuan et Deren,2004] Shiyuan H., Deren L., Vector cellular automata based geographical entity, in GeoInformatics 2004, Proceedings of the 12th Conference on GeoInformatics, University of Gävle, Sweden, 7-9 June 2004, pp. 249-256.

[Sibony, 1991] Sibony D., *Entre-deux, l'origine en partage*, Seuil, collection la couleur des idées, 1991, 400p.

[Sinton 1978] Sinton D., The inherent structure of information as a contraint to analysis : mapped thematic data as a case study, In *Harvard Papers on GIS*, Vo.7, edited G. Dutton, Addison-Wesley.

[Spéry, 1999] Spéry L., *Historique et mise à jour de données géographiques : application au cadastre français*, Thèse de doctorat ès Géographie, Université d'Avignon et des Pays de Vaucluse, 1999, pp. 133-145.

[Thomas, 1996] Thomas I., Spatial data aggrégation : exploratory analysis of road accidents , *Accident Analysis and Prevention*, vol.28, n°2, 1996,p.251-264.

[Tellez et Servigne, 1997] Tellez B., Servigne S., Updating Urban Database with aerial Photographs : A common Structuring Methodology. *Computers, Environment and Urban Systems*, vol.21, n°2, mars 1997, pp. 133-145.

[Timpf, 1997] Timpf S. Cartographic objects in a multi-scale data structure, In *Geographic Information Research* : Bridging the Atlantic, Edited by M. Craglia and H. Couclelis, London, Taylor et Francis, 1997.

[Weber et Hirsch, 2000] Weber C., Hirsch J., Processus de croissance et limite urbaine, Cybergéo 5/5/2000

[Worboys, 1998] Worboys M., A generic Model for Spatio-Bitemporal Geographic Information, In Egenhofer M-J., Golledge R-G., (Eds) *Spatial and temporal reasoning in geographic information systems*, Oxford University Press, 1998, 276p.

[Worboys, 1995] Worboys M.F., *GIS : A computing perspective*, Taylor and Francis, 1995.

[Yang, Lo, 2003] Yang X., Lo C.P., Modelling urban growth and landscape changes in the Atlanta metropolitan area, *International journal of geographical information Science*, Vol. 17, N°5, July-August 2003, p463-488, Taylor and Francis.

[Zaninetti, 2005] Zaninetti J-M., *Statistique spatiale*, Hermès Lavoisier, 2005, 320p.

autres références

Sites Internet

Dictionnaire de l'informatique et d'internet : Dico.fr.
www.dicofr.com/cgi-bin/n.pl/dicofr/definition/20010101002600
consulté le 20/8/2006,

Le serveur éducatif de l'Education National sur l'Information Géographique-
2. Les différents types d'information géographiques
http://seig.ensg.ign.fr/fichchap.php3?NOCONT=CONT3&NOCHEM=CHEMS002&NOFICHE=FP7&NOLISTE=3&N=1&RPHP=&RCO=&RCH=&RF=&RPF=
consulté le 11/10/2006

Système d'information géographique
http://fr.wikipedia.org/wiki/Syst%C3%A8me_d%27Information_G%C3%A9ographique
consulté le 15/10/2006

L'information géographique c'est…
http://www.quebecgeographique.gouv.qc.ca/decouvrir/introduction/introduction.asp
consulté le 18/10/2006

NHGIS (National Historical Geographic Information System)
http://www.pop.umn.edu/about/brochures/nhgis.pdf
consulté le 17/10/2006

Les observatoires du territoire
http://www.ieti.fr/obsterri_accueil.htm
consulté le 6/6/2006

Introduction à l'informatique décisionnelle
http://www.commentcamarche.net/entreprise/business-intelligence.php3
consulté 10/10/2006

EDP et évolution de forme 2D
Yoann Juge et Nicolas Rougon. 1998 :
http://www-eph.int-evry.fr/~horain/atelvir/Segmentation/EvolutionContours/index.html
consulté 4/7/2006

L'océan sous toutes ses facettes
www.jason.oceanobs.com/html/kiosque/dossiers/alti/facettes_fr.html
consulté le 18/10/2006

Géosimulation
hypergeo.free.fr/article.php3?id_article=311
consulté le 18/10/2006

Analyse spatiale, généralités
http://hypergeo.free.fr/article.php3?id_article=38
consulté le 18/10/2006

De l'ordre dans le désordre
http://écrits-vains.com/doxa/tri01.htm
consulté le 15/10/2006

Les scientifiques et les prophètes
http://www.larevuedesressources.org/article.php3?id_article=465
consulté le 18/10/2006

Yann Artus Bertrand, La France vue du ciel
http://www.yannarthusbertrand.com/yann2/affichage.php?reference=TVDC%20FR_27&pais=Francevueduciel
consulté le 24/9/2006

PLU le projet
www.lacub.com/projets/plu/Fichiers/rapport_de_presentation/PLU_B3_1_arrete_07-01-05_zone_u_42a65.pdf
consulté le 19/10/2006

Variation de l'imperméabilisation entre 1982 et 1999
environnement-transport.hauts-de-seine.net/Ressources/pdf/impermeabilisation_17986.pdf
consulté le 18/10/2006

Entretiens évoqués

{AUE, 2006} Etudiants du corps des Architectes et Urbanistes de l'Etat, Interventions lors du cours de SIG, Ecole Nationale des Ponts et Chaussées, Marne-la-Vallée, 2005-2006

{Bonin, 2006} Responsable de l'action de recherche « risques et informations géographiques » au Cogit, IGN, discussion, IGN, Saint Mandé 2006

{Ducuing, 2006} Responsable à l'IGN du partenariat avec les SDIS (Services Départemental d'Incendie et de Secours), Entretien téléphonique, IGN, 27/6/2006.

{Durand-Dastès, 2006} Professeur émérite, Université Paris 7, Entretiens au laboratoire Géographie-Cité, Paris, 2006

{Esposito, 2006}Doctorante, Présentation de son sujet de recherche, Comité scientifique pour les demandes d'allocation de bourse de thèse, Ecole Nationale des Ponts et Chaussée, Marne-la-Vallée, juin 2006.

{MacMaster, 2006}Professor and Chair, Department Geography, University Minnesota, Présentation orale au Laboratoire Cogit, IGN, Saint Mandé, Juin 2006

{Mustière, 2006} Responsable de l'action de recherche « intégration de bases de données géographiques » au Cogit, IGN, Interview, IGN, Saint Mandé, 2006

{Boldo, 2005} Chef du Laboratoire Matis, IGN, Interview, IGN, Saint Mandé, 2005

{Garcez, 2005} Chef du bureau « Stratégies Nationales » à la DGUHC, Ministère de l'Equipement, des Transports, de l'Aménagement du Territoire, du Tourisme et de la Mer, Entretiens téléphoniques en septembre-octobre 2005

{AUE, 2004}Etudiants du corps des Architectes et Urbanistes de l'Etat, Interventions lors du cours de SIG, Ecole Nationale des Ponts et Chaussées, Marne-la-Vallée, 2004-2005

{Foulard, 2004}Géomaticienne, en charge du MOS à l'IAURIF, Entretiens, Ecole Nationale des Ponts et Chaussées, Marne-la-Vallée et IAURIF, Paris, 2002-2004

{Rigail, 2004} DREIF/DUSD/Mission cartographie, Entretien téléphonique, septembre 2004.

{Monjal, 2003}Directeur d'Alphaville (société spécialisée en projets d'aménagement), Entretien, Bureaux d'alphaville, Paris, juin 2003 .

{Morand, 2003}Maitre de conférence à l'université de Marne la Vallées, Cours sur « l'aide à la décision en aménagement urbain » université de Marne la Vallée, mai 2003.

{Sillard, 2003} Direction générale de l'INSEE, Entretien, IGN, Saint-Mandé,

Annexes

Plan des annexes

Annexe 1 : Résolution, précision et granularité : présentation comparée (p. 252)

Annexe 2 : Le *thème*, une approche technique (p. 254)

Annexe 3 : Définition de l'*entre-deux* (p. 257)

Annexe 4 : Une représentation des 'objets' sous forme d'information géographique par nature pour une observation simplifiée dans les SIG (p. 258)

Annexe 5 : Discussions sur l'hypothèse d''emprise au sol' constante (p. 260)

Annexe 6 : Formulation mathématique des indicateurs produits avec la méthode d'observation multi-niveaux (p. 262)

Annexe 7 : Eléments d'applications sur les cas de phénomènes retenus (p. 265)

Annexe 8 : Etapes de mise en œuvre d'une méthode d'observation multi-niveaux (p. 279)

Annexe 1 : Résolution, précision, granularité : présentation comparée

La présentation comparée des trois caractéristiques d'une base de données : *résolution*, *précision* et *granularité*, décline sur l'exemple d'un 'pavillon' les définitions reprises de [Ruas, 2002] (qui regroupe ces notions sous l'appellation de *« niveaux de détails géométriques d'une base de données »)*. Puis nous les présentons dans le contexte pratique de la BD TOPO® de l'IGN.

Soit un objet *pavillon*, soit sa représentation en information géographique 'pavillon' et la composante géométrique d'une représentation de ce 'pavillon' sous forme de donnée (cf. figure a.1.1)

Objet *pavillon* | Objet 'pavillon' | Composante géométrique de la données représentant le 'pavillon'

Figure a.1.1. *Pavillon : objet, 'objet' et composante géométrique d'une représentation en donnée*

On a alors :

- *« la **résolution** est la plus petite unité mesurable »* : elle spécifie l'écart minimal entre deux éléments de représentation. C'est la taille du pixel en mode maillé, la distance minimale entre deux points d'une ligne ou d'un polygone en mode vecteur (cf. figure a.1.2). Elle qualifie en terme technique la *représentation* (comme action) ; Elle concerne la production et la numérisation des données géographiques.

Par exemple la résolution de la BD TOPO® est de 10 cm, car le logiciel de production utilisant un codage sur des entiers, applique systématiquement un arrondi sur les coordonnées {x;y} pris égal à 10 centimètres. La résolution est fixée ici par le logiciel SIG.

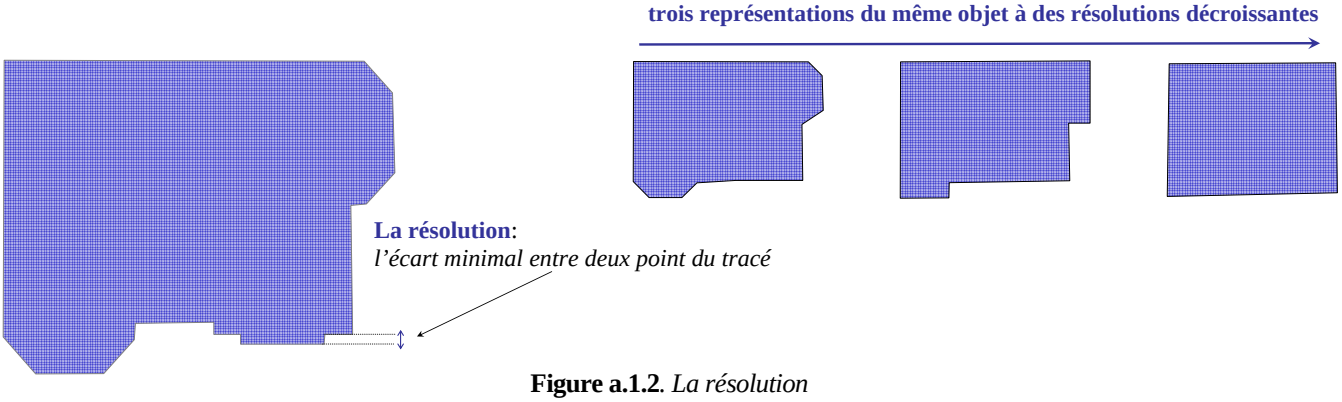

trois représentations du même objet à des résolutions décroissantes
La résolution:
l'écart minimal entre deux point du tracé

Figure a.1.2. *La résolution*

- *« la **précision géométrique** qualifie l'écart entre la position réelle et la position représentée »*. Elle renseigne sur la qualité du positionnement des objets géographiques traduits en données. La précision géométrique peut être appréhendée de façons différentes (cf. figure a.1.3), mais elle concerne toujours les données en tant que produit. Elle qualifie la qualité de la r*eprésentation* (comme résultat)

Par exemple la précision absolue (le meilleur écart que l'on puisse garantir entre un point de la base et la position réelle) de la BD TOPO® est de 1 m. La précision dépend du matériel de saisie.

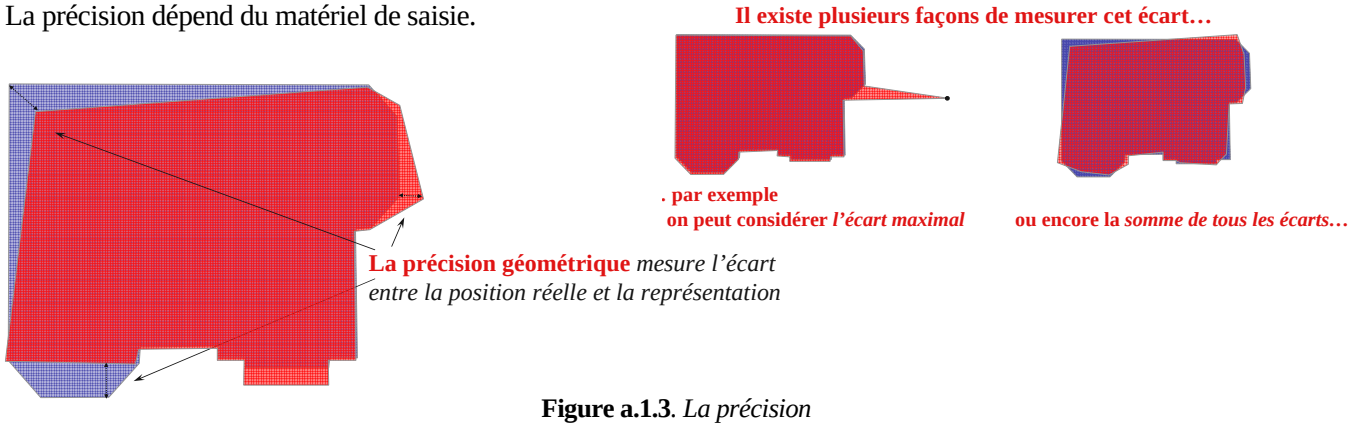

Il existe plusieurs façons de mesurer cet écart…
. par exemple
on peut considérer l'écart maximal
ou encore la somme de tous les écarts…
La précision géométrique mesure l'écart
entre la position réelle et la représentation

Figure a.1.3. *La précision*

- *« la **granularité** décrit la taille des plus petites formes géométriques décrites dans la base.[...][par exemple] la taille minimale d'un décrochement pour un bâtiment »*. Elle dépend de la résolution, mais ne lui est pas synonyme (cf. figure a.1.4).

Par exemple dans la BD TOPO® les décrochements de moins de 2 m ne sont pas saisis pour représenter les bâtiments. La granularité est fixée par les spécifications de saisie.

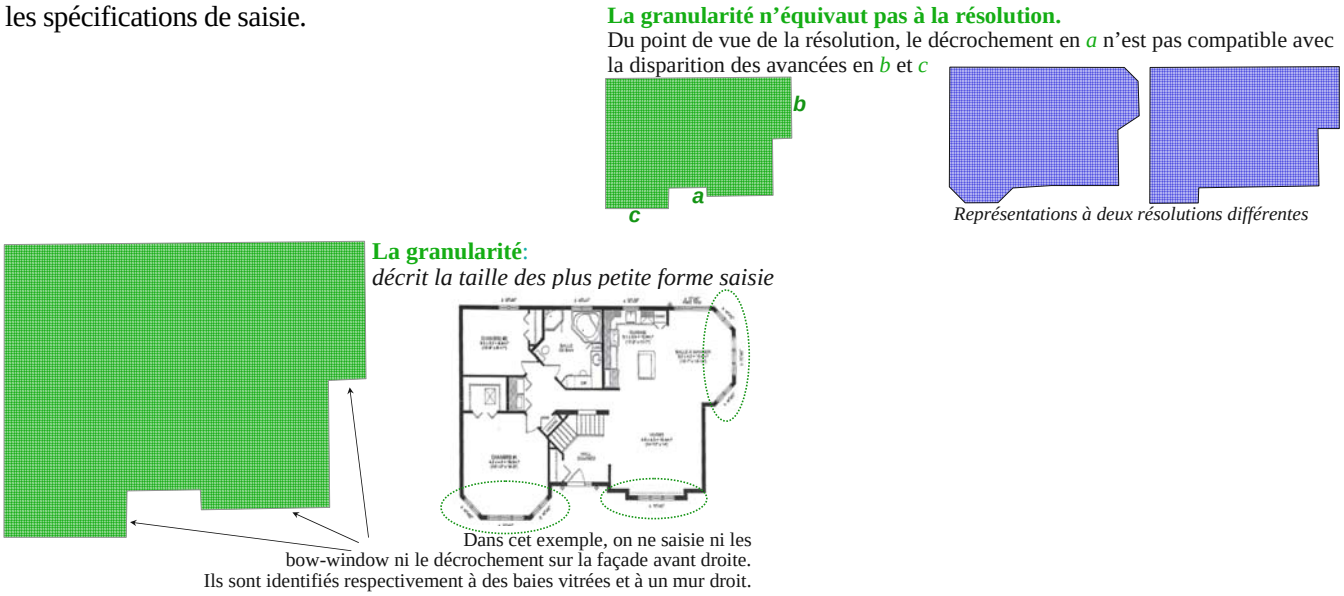

La granularité n'équivaut pas à la résolution.
Du point de vue de la résolution, le décrochement en a n'est pas compatible avec la disparition des avancées en b et c
b
a
c
Représentations à deux résolutions différentes
La granularité:
décrit la taille des plus petite forme saisie
Dans cet exemple, on ne saisie ni les bow-window ni le décrochement sur la façade avant droite.
Ils sont identifiés respectivement à des baies vitrées et à un mur droit.

Figure a.1.4. *La granularité*

Les niveaux de détails ainsi définis intéressent plus les données géographiques (pour décrire la qualité de leur composante géométrique), que les informations géographiques.

Annexe 2 : Les *thèmes :* une approche technique pour spécifier les informations géographiques

Les producteurs de bases de données semblent concevoir la *sélection* des informations par grand *thème* générique donnant à voir des aspects complémentaires du territoire. La BD TOPO® de l'IGN par exemple propose différents « thèmes » d'informations (cf. figure a.2.1 et figure a.2.2).

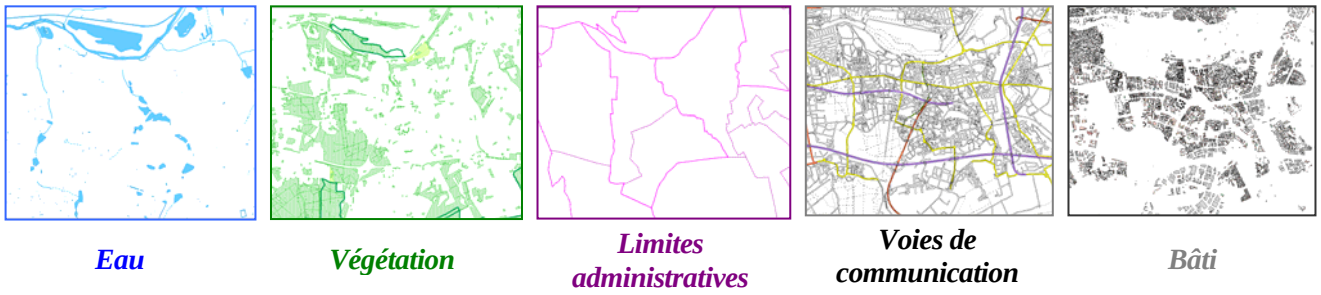


Figure a.2.1. (rappel de la figure B.1.4) *Les grands thèmes de la BDTopo®*

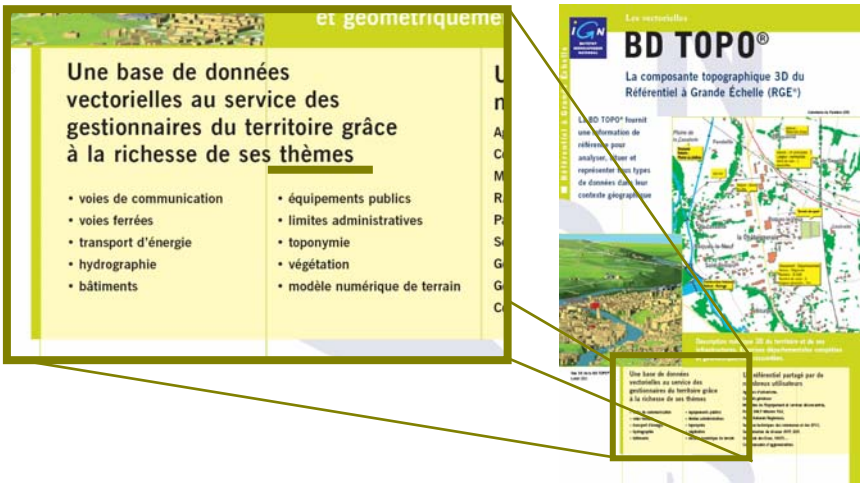

Une base de données vectorielles au service des gestionnaires du territoire grâce à la richesse de ses thèmes
• voies de communication
• voies ferrées
• transport d'énergie
• hydrographie
• bâtiments
• équipements publics
• limites administratives
• toponymie
• végétation
• modèle numérique de terrain
BD TOPO®
La composante topographique 3D du Référentiel à Grande Échelle (RGE®)

Figure a.2.2. *Les grands thèmes de la BDTopo®*

Cette façon de concevoir l'étape de sélection est différente de celles des utilisateurs thématiciens. La conception des *thèmes* et cette approche de la sélection résultent, à notre avis, à la fois de raisons :

- **historiques**,
- **techniques** et
- **pratiques**

se renforçant les unes les autres.

De leurs **origines,** les SIG ont hérité une certaine conception liée aux outils de DAO/CAO (dessin / conception assistés par ordinateur). Les *thèmes* pourraient ainsi être la conséquence des «calques» qui existent encore dans le logiciel de CAO Autocad. Les thèmes facilitent le transfert de données des outils de dessin vers les SIG et donc la récupération de données existantes. Pour d'autres raisons **techniques**, les calques ont trouvé un relais naturel dans les SIG : les « couches ». En effet, pour gérer la topologie l'un des premiers logiciels de SIG, ArcInfo, s'appuyait sur la notion de *graphe planaire*. Cette notion impose que toute intersection géométrique entre deux éléments du graphe soit matérialisée par un point. On voit bien l'intérêt pour les routes. En effet, même lorsque le croisement

ne correspond pas à une connexion réelle, il possède généralement un sens thématique (ex. : *pont, tunnel, voie souterraine*). En revanche, il n'est pas utile de maintenir le graphe planaire entre un 'réseau de télécommunication' et les 'chemins de grandes randonnées'. Les « couches » ont donc été introduites pour isoler les 'objets' sans logique topologique entre eux : elles servent à optimiser le calcul et le maintien automatique de la « planairité » du graphe.

Aujourd'hui, d'autres modes de gestion des questions topologiques existent[1], mais on trouve encore cette notion de *couche*. En particulier, le logiciel Mapinfo gère des *couches* qui ressemblent à des calques : elles se superposent selon un ordre qui contribue à l'affichage cartographique (cf. figure a.2.3 et surtout a.2.4). L'affichage cartographique papier participe probablement aussi au maintien de la notion de *couche*. En effet, au moment d'imprimer, on constate qu'à un thème correspond souvent à une « couche couleur » (ex. : le bleu pour le thème hydrographie, le bistre pour les courbes de niveaux, le vert pour la végétation (cf. figure a.2.1).

1. Par exemple, dans le logiciel-SIG Géoconcept, la topologie n'est pas archivée ; elle est recalculée.

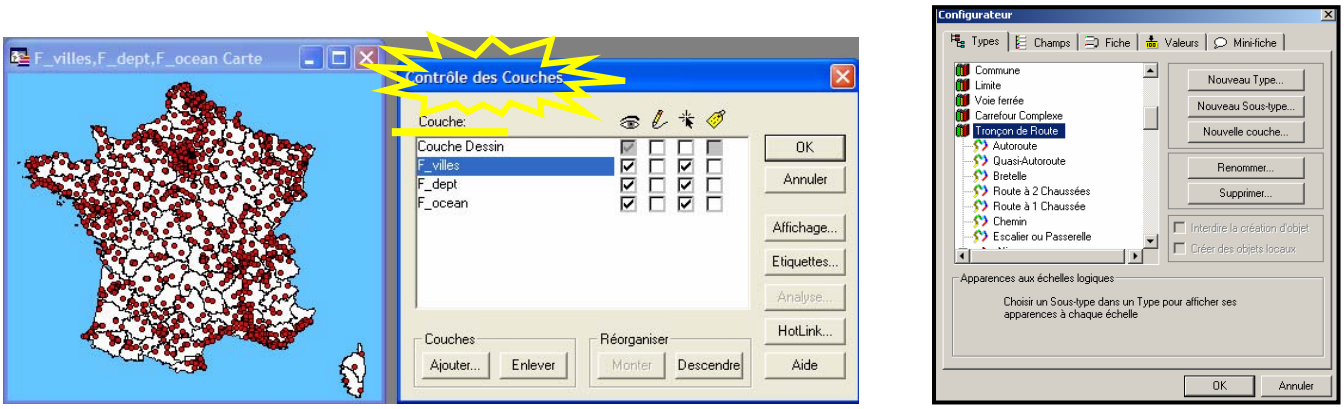


Figure a.2.3. *Différentes façons de gérer les thèmes génériques dans les logiciels de SIG*

Figure a.2.4. *La* couche *dans MapInfo équivalente à un* calque

Cet exemple illustre jusqu'où l'organisation par couche-thème peut devenir technique et « peu » géographique.

*En faisant passer la couche océan au-dessus de la couche des départements (fig a.2.3), on constate que l'organisation des couches dans MapInfo fonctionne comme des calques graphiques. Cet exemple est d'autant plus probant lorsque l'on utilise les données sur les océans fournies en standard avec le logiciel MapInfo : celles-ci ne correspondent pas à des informations géographiques, mais à une couche graphique bleue. En se « souperposant » aux informations géographiques qui décrivent les surfaces terrestres, elle donne à voir par complémentarité la surface maritime. Ainsi sans information sur les pays voisins, la France devient une île (fig. a.2.3) (cela résulte directement de la mauvaise modélisation de l'*océan *en information géographique, et en particulier de son emprise au sol). Cela est particulièrement frappant lorsque l'on fait remonter* l'océan *au-dessus des départements dans l'ordre d'affichage des couches. La France se retrouve alors sous l'eau (fig. a.2.4)*

La notion de *thème* se justifie également par des raisons **pratiques**. Indépendantes de la structuration informatique, elles font échos à la structuration organisationnelle de la production de données. La collecte d'informations diverses nécessite souvent d'aller consulter des documents ou des producteurs différents. Ainsi, les données sur les *voies de chemin de fer* sont obtenues auprès de la SNCF, celles sur les *bassins versants* auprès des agences de bassins, etc. Les *thèmes* peuvent alors correspondre à une structuration des modes d'acquisition et de mise à jour des données. Ces *thèmes* qui restent conçus dans une approche technique s'identifient souvent aux thématiques génériques (ex : les *voies de communication*, *l'hydrographie*, etc.) qui intéressent les utilisateurs. C'est autour de ces *thèmes génériques*, proches du concept de *type* dans Géoconcept (cf. figure a.2.5), que producteurs de données et utilisateurs se retrouvent.

En effet, le *thème* pour les utilisateurs prend un sens différent, lié à des problèmes différents de sélection.

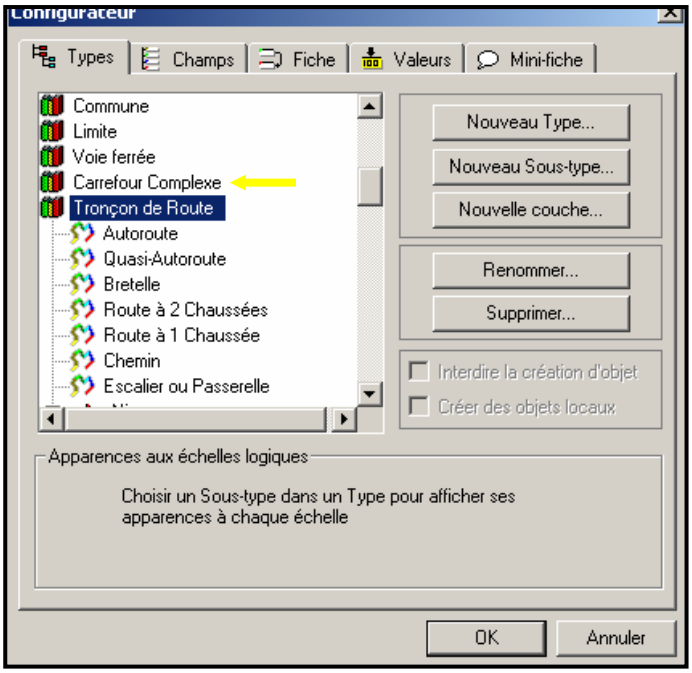


Figure a.2.5. *La notion de type dans Géoconcept proche de la notion de thème pour les utilisateurs*

Sur la figure, on notera la présence de l''objet' de type « carrefour complexe ». Cela montre que les types conçus pour être proches des objets géographiques thématiques (tels que conçus par les thématiciens) peuvent également être utilisés pour prendre en compte des aspects techniques de la modélisation des informations géographiques. En effet, peu d'utilisateurs considèrent les carrefours comme une « thématique », en revanche la complexité topologique de ces objets nécessite un traitement particulier.

Annexe 3 : Définition de l'entre-deux

Le concept *d'entre-deux* posé par [Sibony, 1991] vient élargir la notion de différence : *« La différence se révèle être un cliché de l'entre deux, un cas limite et limité, un cas particulier »*. A. Sibony l'introduit en notant :

« Nous avons vécu et pensé jusqu'ici sous le signe de la différence : différence sexuelle, différence entre autochtones et étrangers ?, …entre mort et vivant (oui, est-il vraiment mort? Peut-on lui prendre ses organes? Pas encore? dans dix minutes…). Il y a toujours eu un trait, une frontière, qui départageait le tout, avec en deçà et au-delà, et qui faisait la différence. Parfois c'était caricatural, « les bons d'un coté les mauvais de l'autre »… L'idée de différence ne suffit plus pour comprendre ce qui se passe… Non que l'idée de différence soit fausse : elle est juste mais limitée, pertinente mais infime… L'entre-deux est une forme de coupure-lien entre deux termes, à ceci près que l'espace de la coupure et celui du lien sont plus vastes qu'on ne le croit ; et que chacune des deux entités a toujours déjà partie liée avec l'autre. Il n'y a pas de no man's land entre les deux, il n'y a pas un seul bord qui départage, il y a deux bords mais qui se touchent ou qui sont tels que des flux circulent entre eux. La différence apparaît comme un entre-deux trop mince, elle coupe là où c'est la coupure même qui ouvre l'espace d'un nouveau lien, elle fixe d'un trait l'écart, là où le vif de l'expérience a lieu au cœur de cet écart qu'elle bouleverse… »

Dans le cadre de la géographie, [Rey, 1996], [Rey, 1998] conceptualise également le terme d'entre-deux. Elle l'utilise pour décrire le territoire de « l'Europe médiane »[2]. *« Cette conceptualisation [lui sert à sortir] l'analyse des notions de situations intermédiaires, d'espace de transition, voire de périphérie… »*. Pour [Rey, 1996] l'entre-deux *« donne la primauté à l'entre, milieu particulier qui incorpore de façon spécifique les propriétés des corps lointains mis en contact »*.

Si nous avons choisi *l'entre-deux* comme positionnement, c'est en partie parce que *« l'entre-deux concerne l'articulation à l'autre »*. Mais surtout parce qu' *« au-delà des recollements que l'entre-deux actualise, là où il prend sa force c'est lorsque, dans son immense foisonnement, il apparaît comme une figure de l'origine »* [Sibony, 1991], c'est à dire pour nous un lieu où l'existant peut être repensé d'une façon neuve, décentré –en sa nouvelle origine- pour mieux intégrer les (pré)requis de chacun. Nous espérons ainsi améliorer le cheminement conduisant des données au suivi de phénomènes géographiques. C'est sans doute aussi à *l'entre-deux* que se trouve l'origine de la *géomatique*[3].

2. *«* L'Europe médiane *est un intitulé de plus en plus utilisé pour nommer un espace qui n'est plus l'Europe de l'Est… et qui ne serait pas l'Europe centrale »* [Rey, 1998].

3. Géomatique : *« Discipline ayant pour objet la gestion des objets à référence spatiale et qui fait appel aux sciences et aux technologies reliées à leur acquisition, à leur stockage, à leur traitement et à leur diffusion »* Ministère des ressources naturelles et de la faune, Québec. {www.mrn.gouv.qc.ca}
ou encore *« l'ensemble des techniques de traitement informatique des données géographiques »* (journal officiel du 14 février 1994).

Annexe 4 : Une représentation des 'objets' sous forme d'information par nature pour une observation simplifiée dans les SIG

L'origine des logiciels SIG à la fois comme outils de représentation cartographique (représentation éminemment statique) et comme outils de gestion de données, les a dotés en standard de fonctionnalités de consultation adaptées aux observations statiques. En effet, en tant que SGBD-S (système de gestion de base de données spatialisées), les SIG sont capables de restituer les données qui leur sont confiées. Ils permettent de répondre aux questions de type *« Où ? » (Où sont les objets ?)* et de type *« Quoi ? » (Quoi qu'il y a là ? Quels sont ces objets ?) ;* Ces questions correspondent à des requêtes élémentaires sur les informations géographiques. Elles interrogent respectivement la composante sémantique et la composante géométrique des données géographiques pour restituer leur localisation (respectivement leurs caractéristiques)[4]. Ces fonctionnalités sont simples à mettre en œuvre et ne posent aucune difficulté technique ; les faits le confirment d'ailleurs puisque dans la pratique les applications de type *inventaire* ou *observatoire* sont très développées. Elles constituent d'ailleurs une « entrée » dans le domaine de la géomatique pour beaucoup d'utilisateurs. Cependant, pour que l'observation d'un 'objet' dans un SIG soit effectivement simple, il est important que l'*objet* possède une représentation explicite avec une composante géométrique propre, autrement dit une représentation sous forme d'*information géographique par nature*. En effet, si l'*objet* ne fait pas partie des informations listées dans les spécifications, les requêtes de type « *Où* sont les 'objets' ?» ne trouveront pas de réponse. Par ailleurs, seule la modélisation sous forme d'information par nature optimise l'exploitation des outils de requêtes, comme ceux présentés dans la figure a.4.1, requêtes géométriques et/ou requêtes sémantiques. Elle permet d'obtenir rapidement les attributs d'une information géographique et/ou sa position, son emprise au sol. Ainsi, la consultation directe par un 'clic' de « souris » (cf. figure a.4.2) nous permet de *saisir* une information géographique sous forme de données: Elle sélectionne sa composante géométrique et l'affiche en surbrillance, elle facilite l'accès à sa composante sémantique. De la sorte, nous disposons de moyens pour observer l'information et ainsi mieux la *saisir* au sens figuré de « comprendre ». Lorsque que nous employons l'expression : « *saisir* … dans un SIG », nous donnons d'ailleurs au mot *saisir* cette double acception (sélectionner et comprendre)[5].

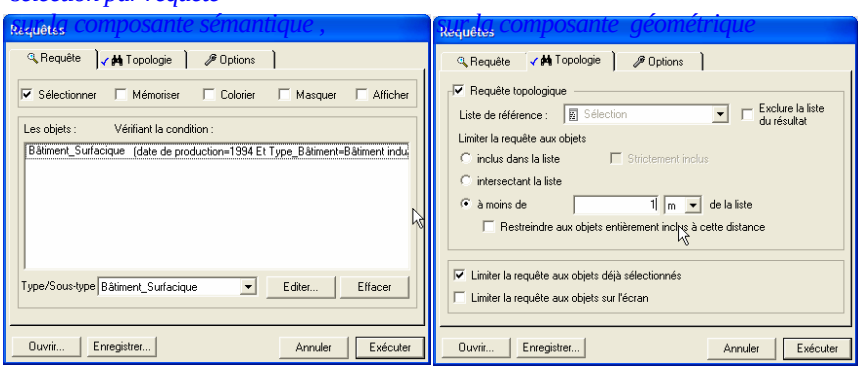
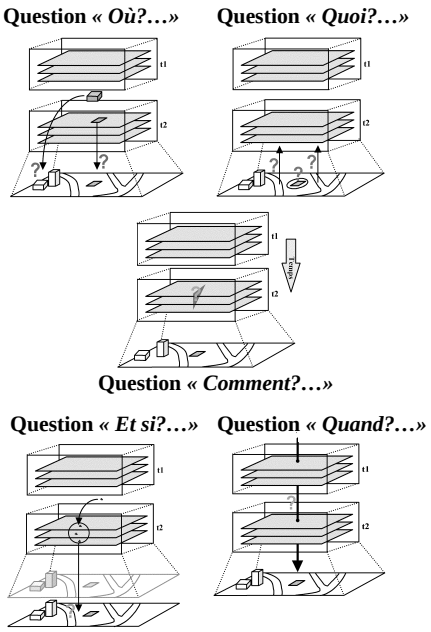


Figure a.4.1. *Exemples d'outils de requêtes élémentaires*

La représentation sous forme d'information géographique par nature permet d'obtenir directement des informations descriptives sur un 'objet ' et sur sa localisation

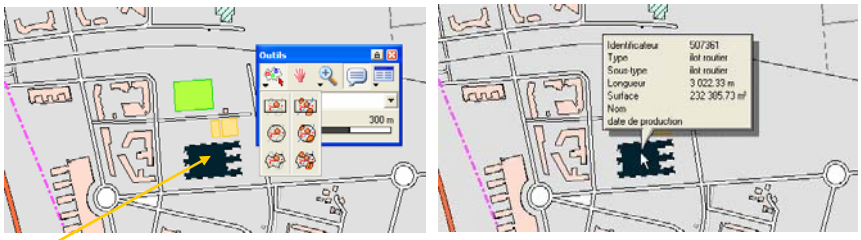
4. Parmi les 5 types de questions posées aux informations géographiques, identifiés par [Denègre et Salgé, 1996] : Où ?, Quoi ?, Comment ?, Et si ? et Quand ?, les questions de type Où ? et Quoi/Quel(lles) ? correspondent au niveau d'analyse dans un SIG le plus simple, celui des requêtes élémentaires. Les questions Comment ?, Et si ?, qui visent à créer une information nouvelle (le plus court chemin, une zone d'influence, les zones inondées en cas de crue, etc.) nécessitent des outils plus élaborés d'analyse à référence spatiale. Enfin, le Quand ? renvoie aux requêtes sur la dimension temporelle



5. Cette expression évoque à la fois un projet de connaissance sur des objets réels et la démarche technique utilisée par ce projet : la consultation de données. Nous n'emploierons pas cette expression pour évoquer l'observation des données dans un SIG dans le cadre d'une application technique (par exemple une mise à jour ou d'un contrôle qualité). Dans ce contexte informatique, l'expression « saisir »… dans un SIG, signifie d'ailleurs habituellement *digitaliser, numériser des données*.

* Le ‘clic’ souris *saisit* la composante géométrique de l’objet sélectionné et l’affiche en « surbrillance » (a)

* Il permet d’accéder simplement à sa composante sémantique soit directement à la « fiche » soit sous forme de bulle (b).



… sélection de la composante géométrique
par sélection directe à l’écran

Figure a.4.1. *Le ‘ clic’ à la souris comme outil de requêtes élémentaires dans les SIG pour observer les ‘objets géographiques’*

Annexe 5 : Eléments de discussion sur l'hypothèse d''emprise constante'

L'observation multi-niveaux dans la classification de [Sinton, 1994]

Partant du constat que les applications sur les phénomènes géographiques peinent à intégrer les trois dimensions (temporelle, géométrique et sémantique), [Sinton, 1978] propose de les classer en fonction du statut octroyé à chacune, en déterminant laquelle est :

– fixée à valeur constante,
– « contrôlée » et prend un nombre limité de valeurs,
– mesurée sur des intervalles.

Reprenant cette approche évoquée dans le paragraphe A.2.4, p.63, nous proposons de situer les applications de la méthode multi-niveaux dans cette classification. Dans une observation multi-niveaux :

- la dimension fixée est : la dimension géométrique (hypothèse de l'emprise constante) ;

- la dimension contrôlée est : le temps. (Les valeurs sont contraintes par les observations, par exemple, avec un modèle par *snapshots* les valeurs correspondent aux dates des mises à jour).

- la dimension mesurée est : la sémantique. (Les caractéristiques sont données sous forme de fonction mathématique).

Toutefois, tout en restant dans le cadre d'une observation multi-niveaux, il est possible d'assouplir les contraintes posées sur les deux dimensions : *temps* et *espace*, pour permettre de travailler sur une dimension géométrique « un peu moins fixe » et sur une dimension temporelle « un peu moins contrôlée ».

Une dimension temporelle un peu moins contrôlée

La dimension temporelle est « contrôlée » au travers des dates d'observation. Dans le paragraphe B.3.2.1.3, nous montrons que l'on peut « assouplir le caractère contrôlé » au sein de la méthode multi-niveaux. En effet, en introduisant des fonctions d'interpolation (et d'extrapolation) sur les valeurs de chaque caractéristique obtenue à partir des versions v_t d'une base de données micro, on peut envisager théoriquement disposer d'une représentation de l'état de l'objet' à n'importe quelle date.

Cela revient à prendre la dimension spatiale comme constante, et à faire varier la dimension sémantique en fonction du temps.

On peut également envisager que l''emprise' ne soit pas constante ; cela revient alors à considérer que les deux dimensions spatiale *et* sémantique sont variables. Pratiquement, cela ne nous permet plus de comparer de la même façon les valeurs prises par les indicateurs d'un état à l'autre (par exemple dans le cas d'un phénomène d'épidémie, on pourra étudier la portion de territoire touchée, celle–ci évoluant, on pourra suivre l'indicateur 'nombre de foyers touchés', pour mesurer l'étendu du fléau. En revanche on ne pourra pas déterminer quelle est la part de l'augmentation des foyers concernés due à l'extension géographique de l'épidémie et celle due à la densification de l'épidémie sur les territoires déjà atteints).

Une dimension géométrique un peu moins fixe

Le découpage en objets méso intermédiaires proposé comme enrichissement de la méthode, peut apporter des solutions dans certains cas où l'on souhaiterait observer des objets méso géographiques à emprise variable.

Nous proposons ici de distinguer parmi les modifications géométriques :

- *la stabilité*
- *l'extension* : la composante géométrique dans l'état suivant contient strictement celle de l'état précédent
- *la rétraction* : la composante géométrique dans l'état suivant est strictement incluse dans celle de l'état précédent
- *le déplacement* : La forme de la composante géométrique est identique mais la localisation de son centre de gravité est différente
- *la déformation* : la composante géométrique est différente entre les deux états et cela ne correspond ni à une rétraction, ni à une extension, ni a un déplacement.

D'autres typologies auraient pu être proposées plus proche de la problématique de la constance de l'emprise[6]. Celle que nous avons retenue a l'avantage de s'adapter aux conditions d'application de la méthode multi-niveaux. En effet, si la *stabilité* correspond à notre hypothèse initiale (constance de l'emprise), des solutions peuvent également être envisagées pour traiter des cas d'*extension* et de *rétraction* (dans ces deux cas l'hypothèse d'emprise constante n'est plus respectée). La solution que nous envisageons reprend le principe de l'application différenciée de la méthode (cf. paragraphe B.2.3.2) consistant à introduire une partition de l''objet méso' observé en 'objets méso-structurels' plus fins (ex. : la 'zone pavillonnaire' découpée en 'îlots'). Il s'agirait de construire une partition de l''emprise' en deux 'objets méso-structurels' : l'un correspondant à la 'portion de territoire' commune aux deux états du phénomène, l'autre à la partie supplémentaire (dans le cas d'une extension), à la partie retranchée (dans le cas d'une rétraction). Le principe serait assez proche de celui mis en œuvre dans le MOS. Il impliquerait de recoder les caractéristiques sur chaque sous partie. Mais la méthode d'agrégation par emprise simplifierait ce travail. Le développement d'une telle application implique toutefois de pouvoir trouver des 'objets méso' plus fins pour décrire ces parties spécifiques et de définir une stratégie d'étude pour comparer les deux états (par exemple, est-ce que l'on compare l'objet global et on regarde ses variations dues aux parties spécifiques ou bien considère t'on la partie commune et on étudie les évolutions des extensions de façon particulière)…

Concernant les autres cas de modification d'emprise, nous n'avons pas de pistes sérieuses, même si il nous paraît intéressant de noter que dans le cas des déplacements la forme de l'emprise est exactement la même ; seule la localisation change. Théoriquement, on peut envisager d'effectuer la même agrégation sur le même type d''objets micro', en revanche, les objets micros agrégés ne seront pas les mêmes. On pourra alors choisir de comparer les indicateurs sur la base de l'*emprise* commune (et non pas sur la base de la *portion de territoire* commune) (ex. : on pourra comparer le nombre d''objets micro' contenus). En revanche il sera difficile de déterminer dans les évolutions la part liée au temps et celle liée au spécificités géographiques locales.

6. Deux semblent pertinentes ; elles privilégient chacune l'un des deux aspects de la composante géométrique : la forme ou la localisation.

* *Celle qui privilégie la forme* distingue :
- les transformations où la forme reste identique. Il s'agit soit de l'*identité* soit des *déplacements* (au sens mathématique de composition de rotations-translations),
- Les *transformations* où la forme change. Elles incluent les transformations régulières telles que les homothéties (donnant lieu aux rétractions / extensions).

* *Celle qui privilégie la localisation* distingue :
- les transformations qui ne modifient pas la localisation. Il s'agit soit de l'*identité soit* des *déformations,* regroupant les *transformations* précédentes, (c'est à dire les modifications de formes, même régulières comme les homothéties) à la condition qu'elles ne modifient pas le centre de gravité de la forme.
- Les *déplacements* (dans un sens plus large que le sens mathématique) où la localisation est modifiée. On envisagera ici tous les changements modifiant le centre de gravité, qu'ils conservent ou pas la forme.

Annexe 6 : Formulation mathématique des indicateurs produits avec la méthode d'observation multi-niveaux

a.6.1. Principes généraux de la méthode : formulation d'une observation multi-niveaux statique sur un unique objet

Soit *O* l'objet méso géographique observé et *E* son emprise.

Soit $\{o_i\}$ les objets micro contenus dans la base des données micro utilisées pour effectuer l'observation multi-niveaux.

On détermine alors *A*, l'agrégation des objets micro caractérisant *O* en utilisant *E* et la méthode d'agrégation par emprise :

$$A = \{o_i / o_i \in E\}$$

Pour caractériser *A* au sein de *E*, on détermine un certain nombre d'indicateurs sous forme de fonction f_k de *A*, et éventuellement de *E* (par exemple pour la densité du bâti dans une zone). On notera alors *:*

$$f_k(A) = f_k(\{o_i / o_i \in E\} \text{ et de } E)$$

Après identification formelle des deux 'portions de territoire' : celle de *A* et celle de *O*, ces valeurs $f_k(A)$ sont intégrées à la composante sémantique (c.a.d. aux attributs de *O).*

Dans une observation statique cependant, pour observer un phénomène, on est généralement amené à considérer plusieurs objets méso (ex. : observation de l'urbanisation pavillonnaire, implique d'observer plusieurs zones pavillonnaires).

a.6.2. Principe d'une observation statique multi-niveaux : formulation d'une observation multi-niveaux statique sur plusieurs objets

Soit $\{O_j\}$ l'ensemble des objets géographiques à partir desquels on souhaite observer le phénomène. On note E_j leur emprise respective.

Soit $\{o_i\}$ les objets micro contenus dans la base des données micro utilisées pour effectuer l'observation multi-niveaux.

Pour chaque objet O_j, en utilisant son emprise E_j et la méthode d'agrégation par emprise, on détermine alors les A_j, tels que :

$$A_j = \{ o_i / o_i \in E_j \}$$

On calcule alors pour chaque A_j, les valeurs des indicateurs f_k caractéristiques de ces A_j *:*

$$f_k(A_j) = f_k(\{oi / oi \in E_j\} \text{ et de } E_j)$$

Les $f_k(A_j)$ sont ensuite respectivement intégrés aux attributs des O_j sous forme d'attribut que nous noterons : $f_j ; _k$

Sous réserve d'un nombre d'objets O_j observés suffisant, on pourra alors étudier les corrélations entre les $f_j ; _k$. On cherchera à partir des vecteurs ($f_j ; _k$) où k varie, s'il existe des corrélations entre les caractéristiques f_k.

a.6.3. Principe d'une observation multi-niveaux cinématique formulé sur un objet unique

Pour une observation cinématique sur un objet unique, on considère les différents états O^p de cet objet *O* aux différentes dates t^p.

On note E^p l'emprise de O^p, c'est dire l'emprise de *O* à la date t^p.

Par hypothèse nous avons posé que E^p était constante. On la notera donc par E.

Soit $\{o_{i;p}\}$, les objets micro contenus dans la version à la date t^p de la base des données micro utilisées pour effectuer l'observation multi-niveaux.

On définit alors de manière analogue aux cas précédents, les agrégations A^p des objets micros contenus à la date t^p dans l'emprise E de O^p :

$$A^p = \{ o_{i;p} / o_{i;p} \in E \}$$

Contrairement aux E^p, les A^p n'ont pas de raison d'être constantes dans le temps. Ainsi, pour chaque date t^p *et chaque* A^p correspondant, on calculera la valeurs des indicateurs fonctions f_k. On les notera alors :

$$f_k^P = f_k(A^p) = f_k(\{o_{i;p} / o_{i;p} \in E \} \text{ et de } E)$$

Par construction, on peut alors considérer la suite $f_k^* = \{ f_k^P \}$ p variant, comme une série temporelle et étudier pour l'objet O, l'évolution dans le temps de chaque indicateur f_k sur O (ou encore l'évolution des valeurs de f_k prise sur les O^p).

a.6.4. Suivi d'un phénomène géographique à partir de l'observation multi-niveaux cinématique de plusieurs objets : formulation générale

Soit $\{O_j\}$ l'ensemble des objets géographiques à partir desquels on souhaite observer le phénomène. On note E_j leur emprise respective.

Pour une observation cinématique, on considère les différents états O_j^p des objets O_j aux différentes dates t^p. On note E_j^p. L'hypothèse de constance des emprise dans le temps revient à posé $E_j^p = E_j$ quelque soit p.

Soit $\{o_{i;p}\}$, les objets micro contenus dans la version à la date t^p de la base des données micro utilisées pour effectuer l'observation multi-niveaux.

On détermine alors pour toutes les dates t^p, et tous les objets O_j, les agrégations A_j^p des objets micros contenus à la date t^p dans l'emprise E_j de O_j^p :

$$A_j^p = \{ o_{i;p} / o_{i;p} \in E_j \}$$

On calcule alors pour chaque A_j^p, les valeurs des indicateurs $(f_j ; _k)$ à la date t^p, caractéristiques des A_j^p :

$$f^p{}_j ; _k = f_j ; _k (A^p) = f_k(A_j^p) = f_k(\{o_{i;p} / o_{i;p} \in Ej \})$$

Les $f_k(A_j^p)$ correspondant à la valeur de l'indicateur f_k sur les objets micro $(o_{i;p})$ (c.a.d. les objets micro appartenant à la version de la base micro à la date t^p) contenus dans l'emprise de l'objet O_j.

Ces valeur sont ensuite respectivement intégrées aux attributs des O_j^p sous forme d'attribut que nous noterons : $f^p{}_j ; _k$.

Sous réserve d'un nombre d'objets O_j observés et d'un nombre de date t^p correspondant aux versions de la base de données micro qui soient tous deux suffisants, on pourra alors étudier des corrélations entre les $f^p{}_j ; _k$ (cf. fig. a.6.1)

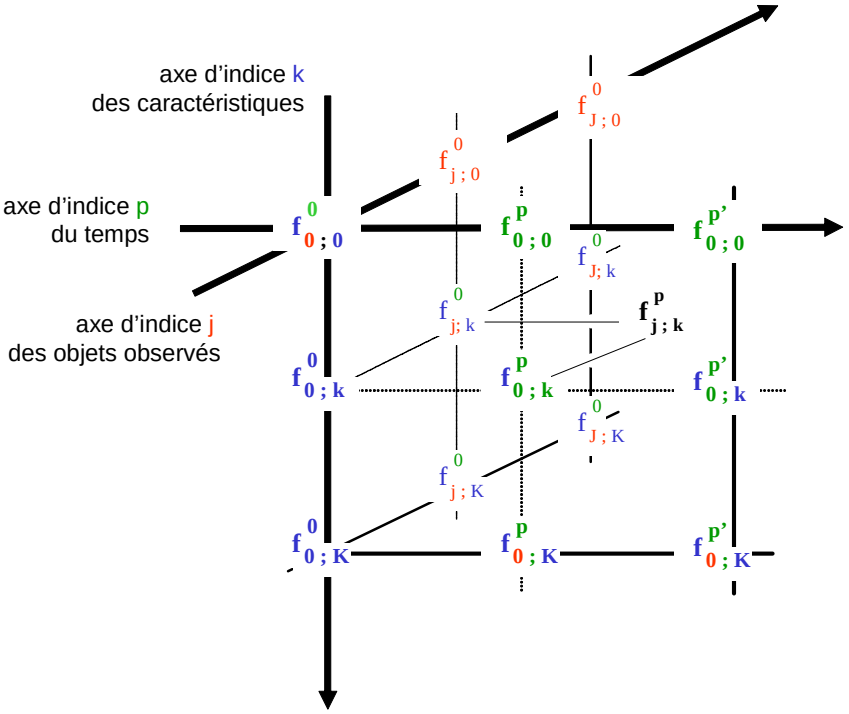


Figure a.6.1. *Dimensions des indicateurs de suivi*

Annexe 7 : Eléments d'application sur les cas de phénomènes retenus

a.7.1. Eléments d'application pour l'observation des zones d'habitat individuel

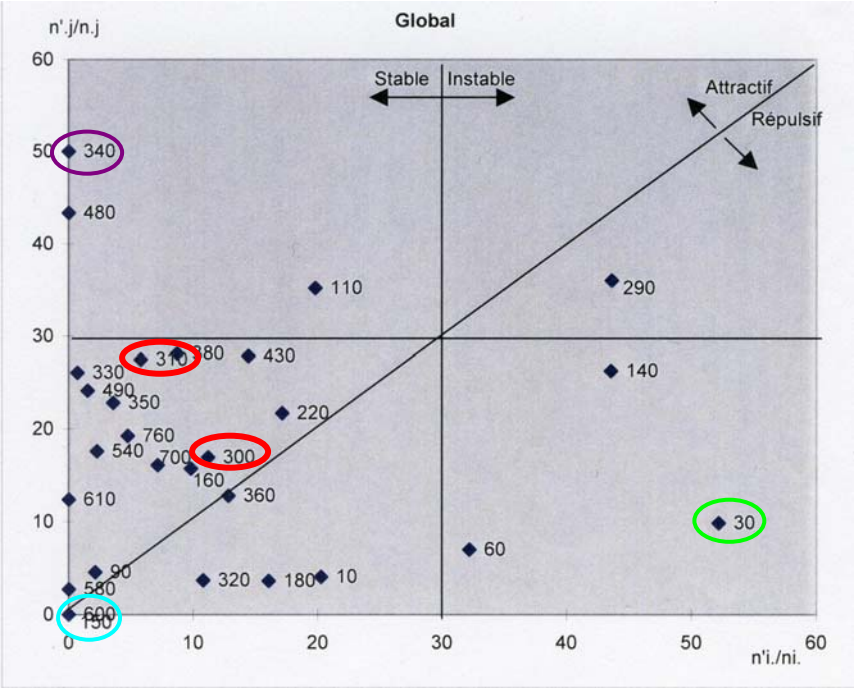


Figure a.7.1. *Eléments d'analyse temporelle des types d'occupation des sols*

Dans le cadre des travaux menés avec Hélène Mathian de l'UMR Géographie-cités, sur l'évolution de l'occupation de l'habitat pavillonnaire à partir du MOS, l'analyse statistique des couples (type d'occupation des sols de la parcelle n à t_i ; type d'occupation des sols de la parcelle n à t_{i+1}) pour toutes les parcelles contenues dans le MOS et sur toutes les dates dont nous disposions (1982 ; 1984 ; 1990 ; 1994 ; 1999) lui a permis à d'établir le graphique de la figure a.7.1. Ce graphique distingue les types d'occupations des sol en fonction d'une part de leur stabilité/instabilité (qui mesure la tendance à rester du même type ou au contraire à se transformer en un type d'occupation des sols différent) et d'autre part de leur attractivité/ « répulsivité » (qui mesure le fait que de nombreux types d'occupation des sols tendent à se transformer en un type particulier, où au contraire tendance d'un type particulier à se transformer en de nombreux autres types) (résultats non publiés).

On remarque alors la très grande stabilité de deux occupations des sols particulières 150 et 600 (les parcelles ayant ce type d'occupation des sols restent avec cette occupation des sols), ni attractives, ni répulsives (c'est à dire aucun autre type d'occupation n'est devenu 150 ou 600) : il s'agit des cimetières (600) et des berges (150).

Le type 30 en revanche est instable et tend à se transformer en différents types d'occupation des sols : il s'agit des surfaces en herbes agricoles, terres labourées, tandis que le 340 est quant à lui plutôt stable et attractif (il peut provenir de divers types d'occupation) : il s'agit de l'habitat collectif continu haut.

Les zones d'habitat individuel sont codées 300 (lorsqu'il s'agit d'habitat individuel) et 310 (pour l'habitat individuel identique). On constate qu'elles sont effectivement stables voire très stables pour les zones d'habitat individuel identique et plutôt attractive.

Parcelles du MOS ayant évolué

Nombre de Bâtiments créés par parcelle du MOS

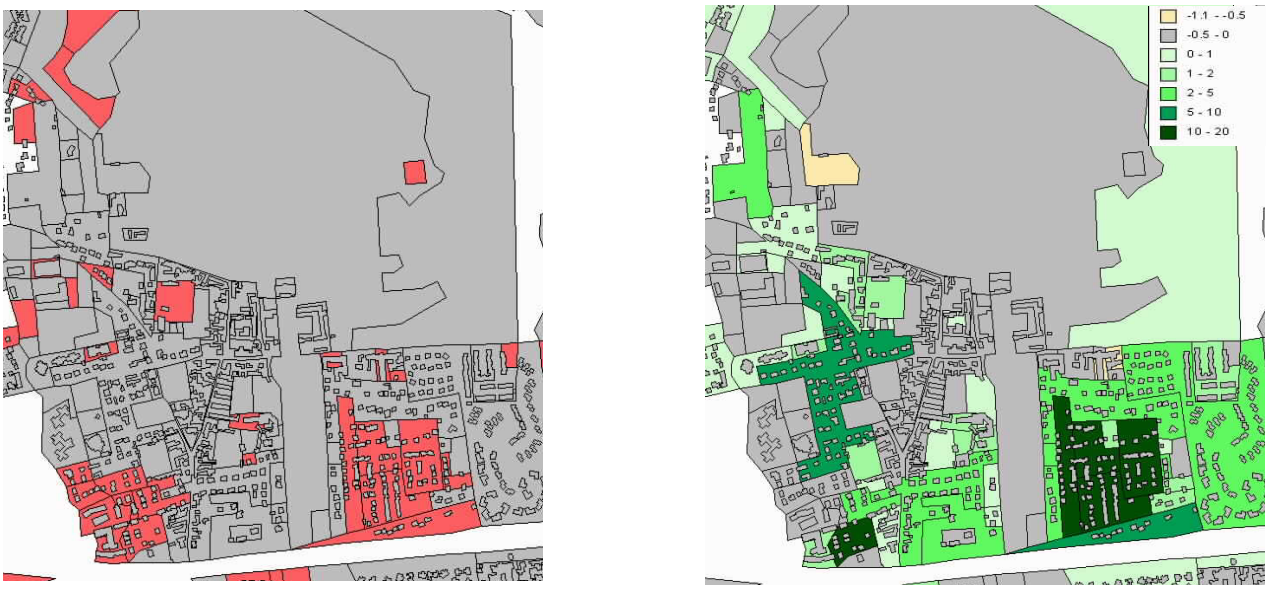


Figure a.7.2. *Comparaison des évolutions dans le MOS avec les évolutions du bâti (création de bâtiments nouveaux) obtenues à partir de la BD TOPO®*

Après avoir intégré le MOS et la BD TOPO® dans un même référentiel, nous avons observé les deux bases conjointement. Les dates de référence pour les mises à jour étant semblables (1994 et 1999), nous avons mis en relation les évolutions saisies dans le MOS et celles que l'on obtenait sur les agrégations d''objets micro' : 'bâtiments' issus de la BD TOPO® (fig. a.7.2). La carte de gauche montre en rouge, les parcelles qui ont évolué dans le MOS sur la période étudiée. La carte de gauche montre sur la même partition en parcelle (même 'objet méso-structurel') les 'parcelles' qui ont évolué en nombre de 'bâtiments'. On constate alors que l'on identifie l'apparition d'une « zone pavillonnaire » ex-nihilo dans les deux cartes ('parcelles' présentées en vert foncé dans la carte de droite). En revanche, la densification de certaines parcelles (classes 2-5) ne s'est pas traduit en évolution dans le MOS et réciproquement des évolutions du type d'occupation des sols ne se sont pas traduites sur le nombre de 'bâtiments'.
(Remarque : le chiffre de -0,5 correspond au fait que le MOS et la BD TOPO® n'appartenant pas au même système de données (cf. p.190), certains 'bâtiments' se trouvent « à cheval » sur deux parcelles) (cf. figure a.7.4)

a. b.

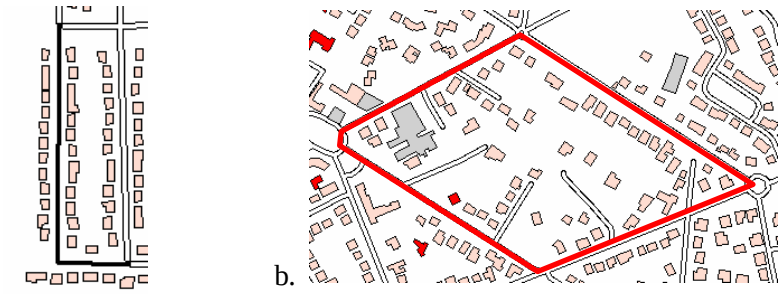


Figure a.7.3. *Utilisation de la caractérisation morphologique*

Dans l'étude réalisée avec Olivier Raimond sur « l'identification automatique des zones pavillonnaires avec des données topographiques », nous avons introduit deux éléments de caractérisation morphologique sur les îlots : la sinuosité des 'tronçons de route' (a) (comme critère de régularité de forme) et des indicateurs pour caractériser la présence d'impasse dans les 'zones pavillonnaires' étudiées (b). Ces indicateurs fournis par Cécile Duchêne sont issus de ses recherches en généralisation [Raimond, 2005] ; [Duchêne, 2004]

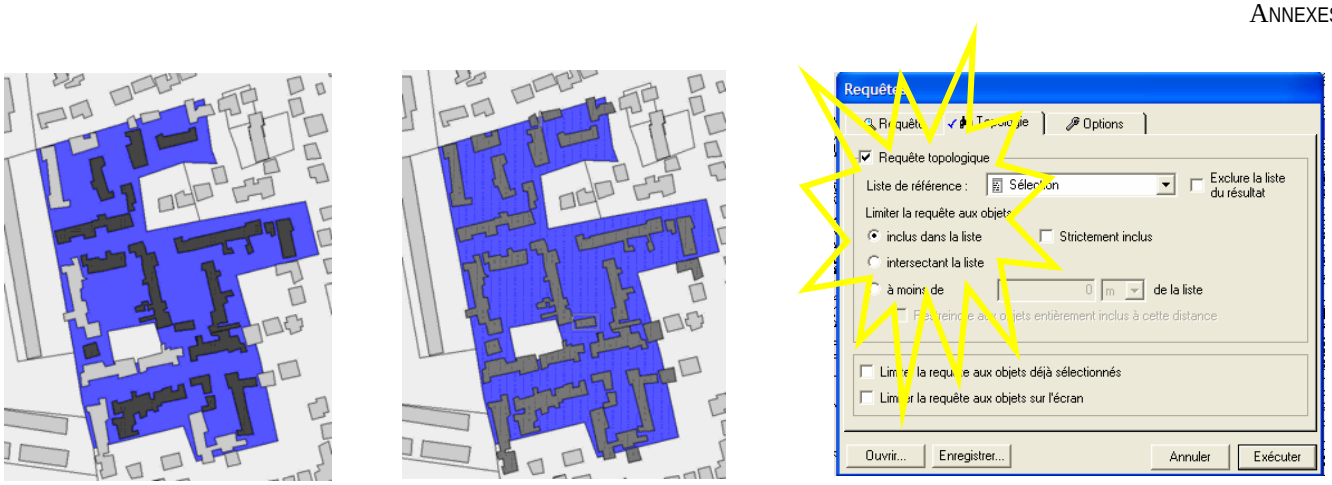


Figure a.7.4. *Exploitation basique de la méthode d'observation*

Lorsque l''objet méso-structurel' n'appartient pas au même système de données que les données micro à agréger, l'utilisateur se trouve alors souvent confronté à des problèmes d'incohérences topologiques. Ce fut le cas, lorsque pour observer les 'zones pavillonnaires' nous avons agrégé le 'bâti' de la BD TOPO à l'aide de 'parcelles d'habitat individuel' du MOS. Dans un premier temps, sans intégrer la notion de « portion de territoire » nous avons choisi de résoudre le problème en définissant deux catégories de caractéristique : les caractéristiques sur l'agrégation stricte (les 'bâtiments' strictement inclus dans l'emprise d'agrégation) et les caractéristiques sur l'agrégation large (les 'bâtiments' inclus ou intersectant l'emprise d'agrégation).

Dans un second temps, nous avons numérisé les 'îlots routiers' à partir de la BD TOPO®, pour pouvoir travailler avec des 'objets méso-structurels' appartenant au même système de données que celui des 'objets micro' 'bâtiments' à agréger. Nous avons alors établi un certain nombre de caractéristiques basiques (nombre de bâtiments, surface moyenne, densité par îlot, etc.). Puis disposant de deux versions de la BD TOPO®, nous avons calculé ces caractéristiques pour chaque 'îlot' pour les deux versions de la base. Nous avons alors constaté qu'il était simple d'extraire les évolutions de ces caractéristiques entre les deux états. Les figures suivantes présentent les résultats pour les caractéristiques : nombre de 'bâtiments' (fig. a.7.5) et surface moyenne du 'bâti' (fig. a.7.6).

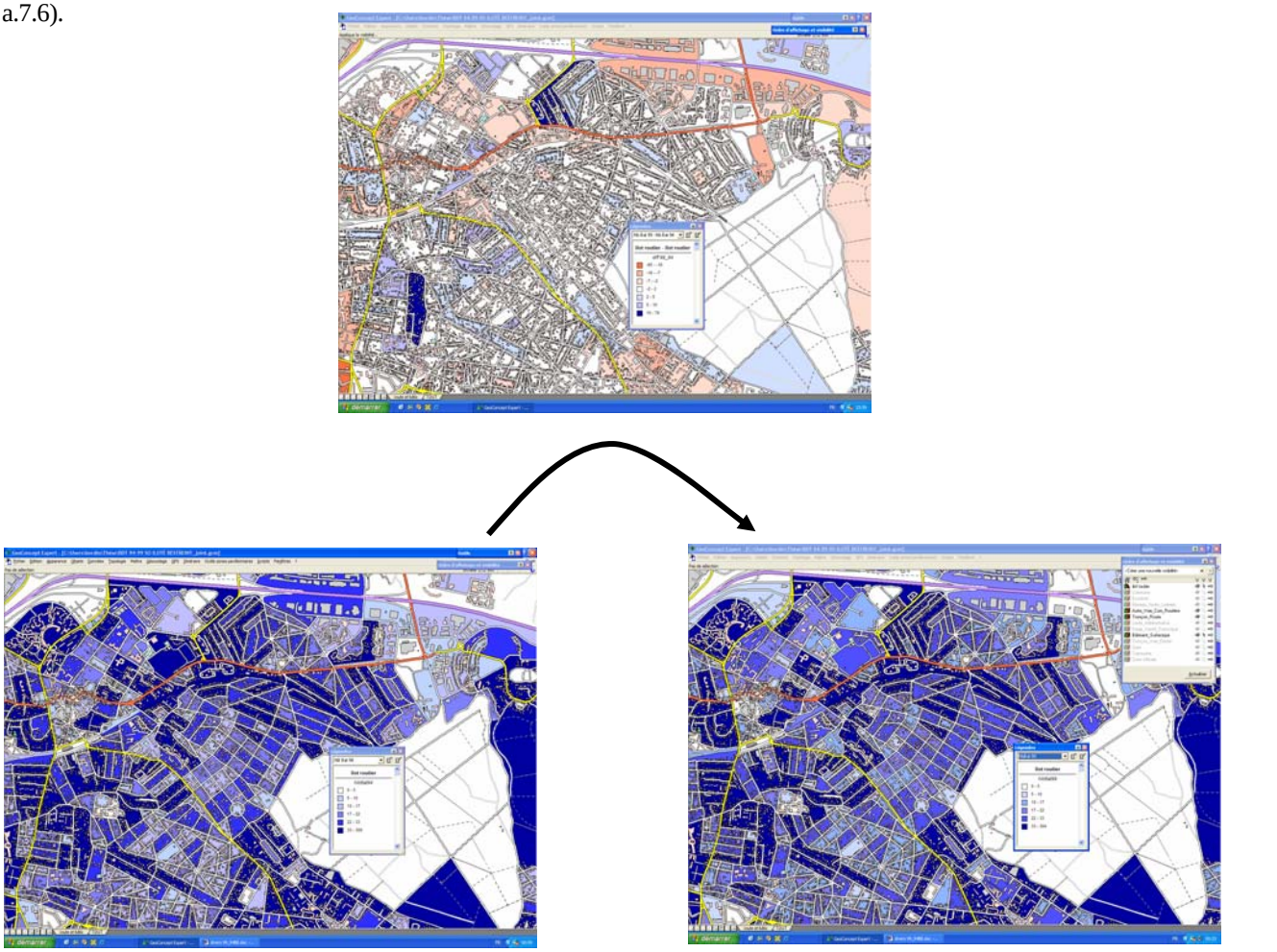


Figure a.7.5. *Evolution du nombre de 'bâtiments' entre 1994-1999 par 'îlot'*

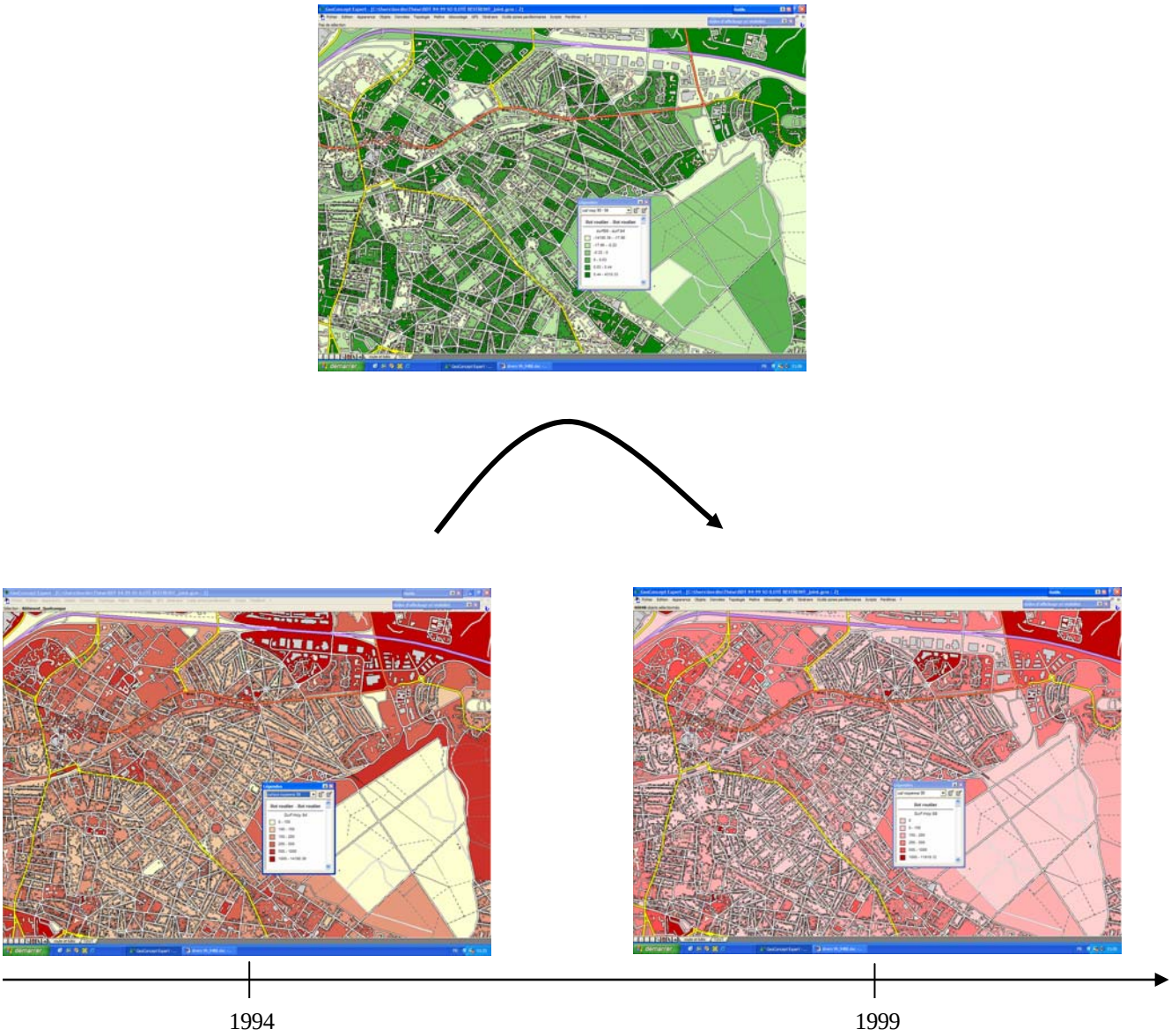


Figure a.7.6. *Evolution de la surface moyenne du 'bâti' par 'îlot' entre 1994-1999*

Exploitation basique de la méthode d'observation (2)

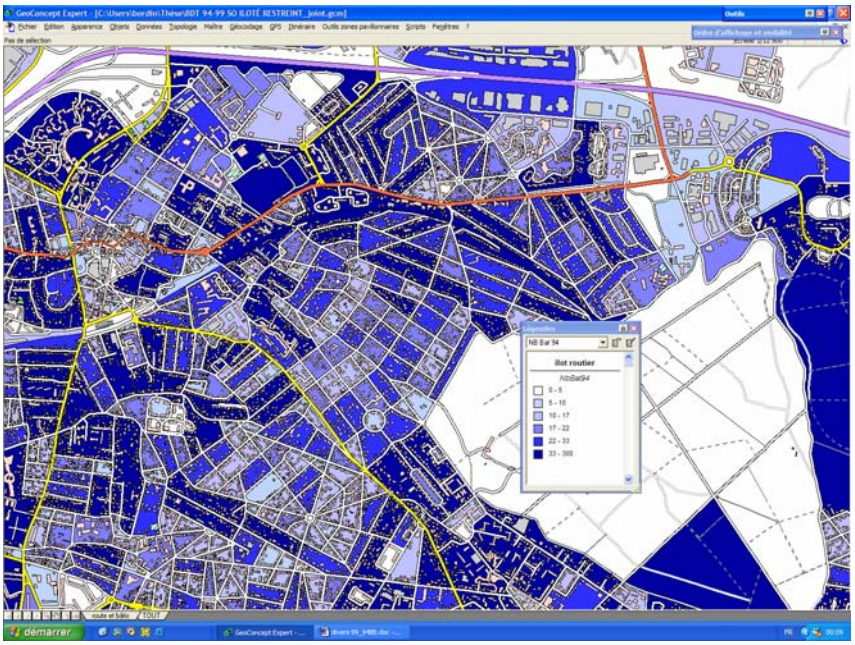
Les six cartes précédentes sont présentées agrandies dans les deux pages suivantes (cf. figure a.7.7). Elles permettent de constater que le fait de pouvoir disposer directement des évolutions permet à l'utilisateur de voir plus facilement et plus rapidement ce qui a changé entre deux états ; cela lui évite une comparaison visuelle longue, fastidieuse, et éventuellement incertaine.

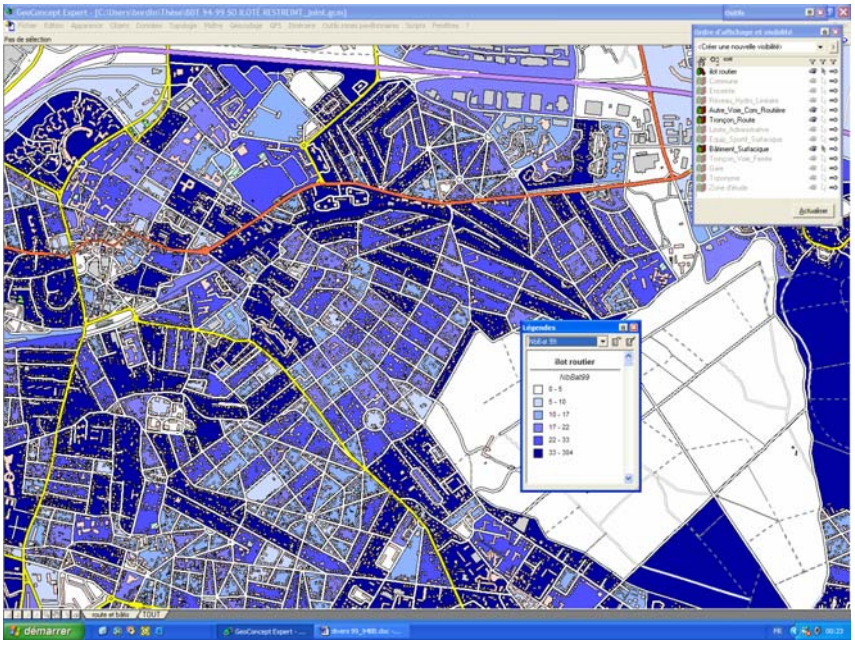
(Remarque : ces cartes obtenues avec des données « vraies » n'ont cependant qu'une valeur illustratives, il ne s'agit pas ici de représenter les cartes des états, mais de montrer les résultats de la mise en œuvre de la méthode d'observation ; cela explique pourquoi nous avons gardé l'«environnement » SIG et pourquoi nous n'avons pas fait apparaître d'échelle)

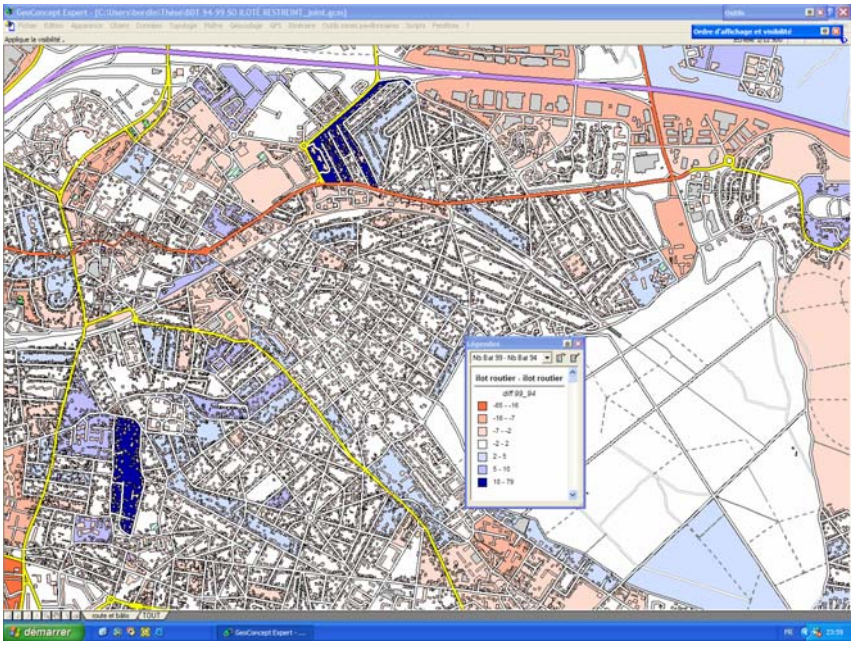
Figure a.7.7. *Deux exemples de cartes d'états en 1994 et en 1999 et cartes des évolutions correspondantes*

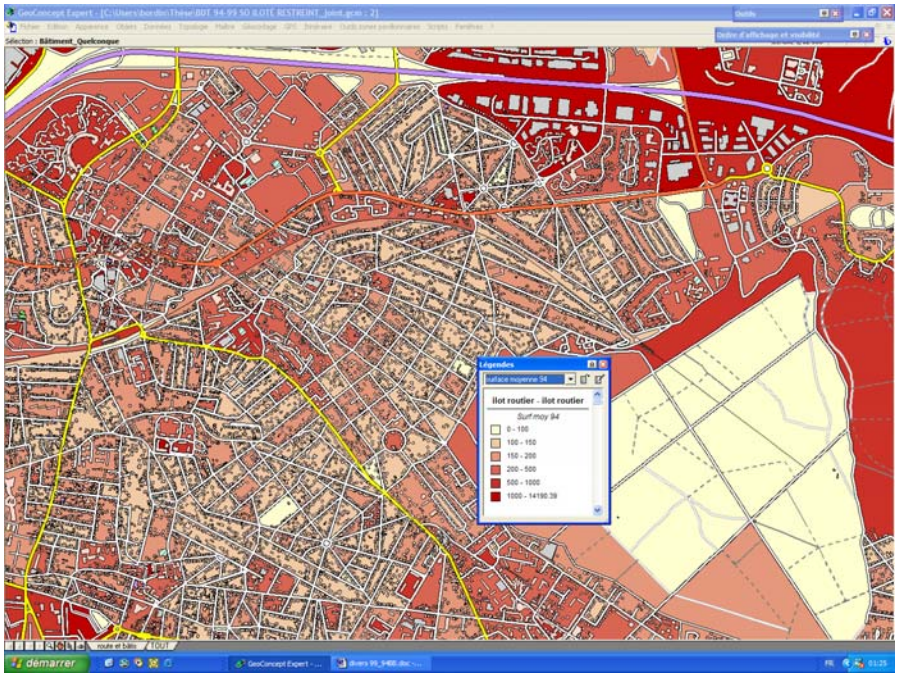
(première page) : Etat des îlots routiers en fonction du nombre de 'bâtiments'

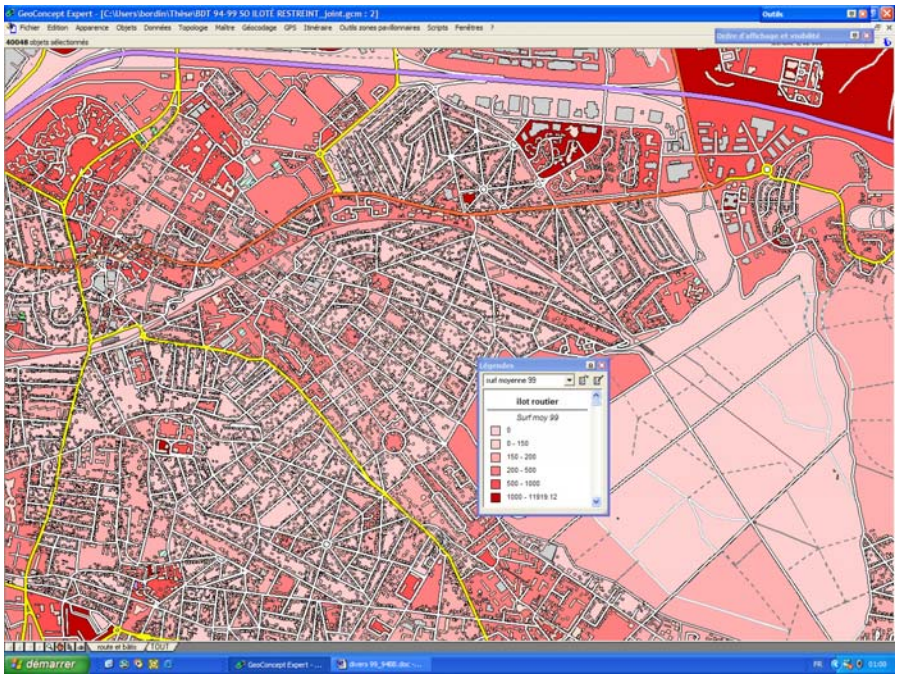
(deuxième page) : Etat des îlots routiers en fonction de la surface moyenne du bâti

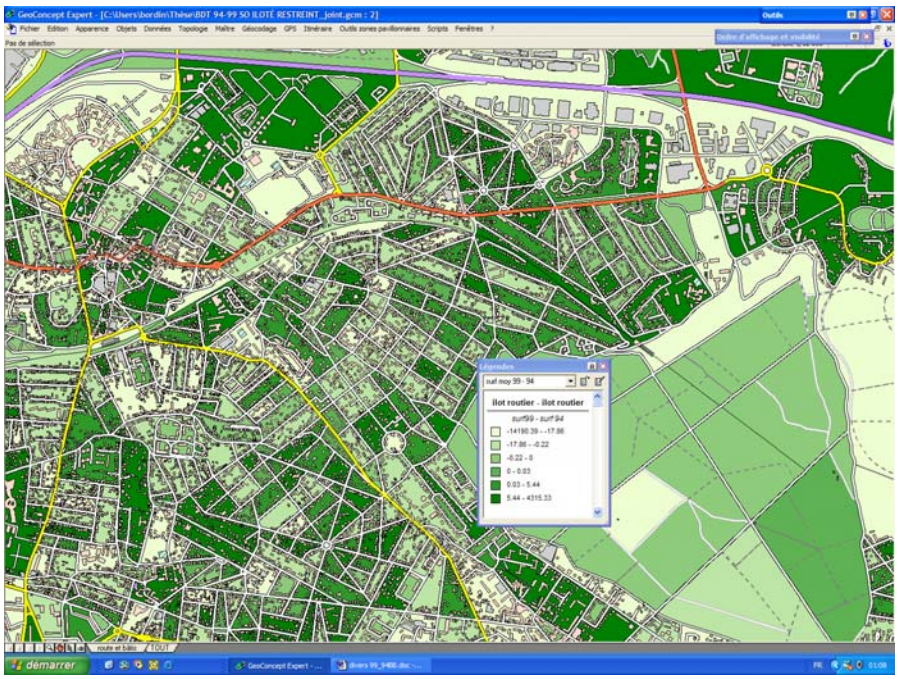

Légendes
NB Bat 94
Ilot routier
NbBat94
0 - 5
5 - 10
10 - 17
17 - 22
22 - 33
33 - 300


Légendes
Ilot routier
NbBat99
0 - 5
5 - 10
10 - 17
17 - 22
22 - 33
33 - 304


Légendes
Nb Bat 99 - Nb Bat 94
Ilot routier - ilot routier
diff 99_94
-65 - -16
-16 - -7
-7 - -2
-2 - 2
2 - 5
5 - 10
10 - 79


Légendes
ilot routier - ilot routier
Surf moy 94
0 - 100
100 - 150
150 - 200
200 - 500
500 - 1000
1000 - 14190.39


Légendes
ilot routier
Surf moy 99
0
0 - 150
150 - 200
200 - 500
500 - 1000
1000 - 11919.12


Légendes
ilot routier - ilot routier
surf99 - surf 94
-14190.39 - -17.86
-17.86 - -0.22
-0.22 - 0
0 - 0.03
0.03 - 5.44
5.44 - 4315.33

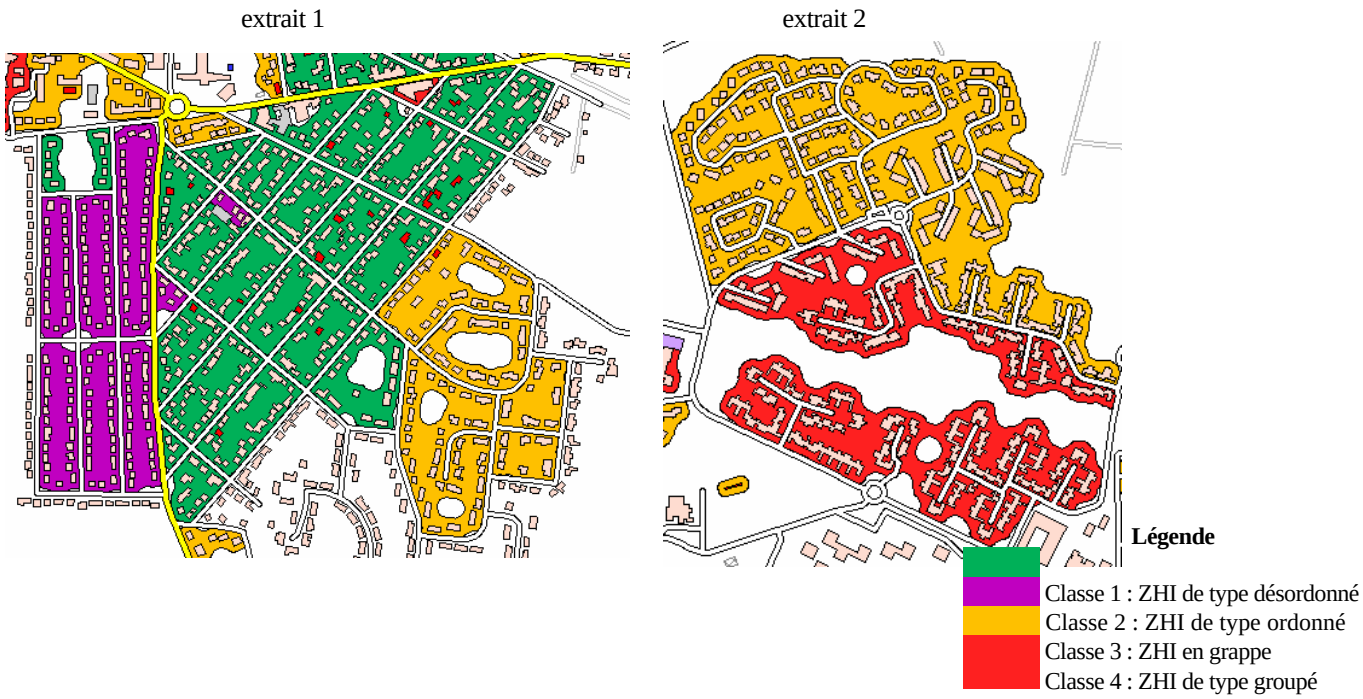

extrait 1
extrait 2
Légende
Classe 1 : ZHI de type désordonné
Classe 2 : ZHI de type ordonné
Classe 3 : ZHI en grappe
Classe 4 : ZHI de type groupé

Figure a.7.8. *Caractérisation des' zones d'habitat individuel' en fonction de la caractérisations des ' îlots' qui les constituent*

Ces deux extraits de [Raimond, 2005] montrent les premiers résultats obtenus par la mise en application de la méthode d'observation multi-niveaux pour l'observation des 'zones pavillonnaires' à partir des 'bâtiments' *et* des 'tronçons de routes' de la BD TOPO® et de l'utilisation des 'îlots routiers' comme objets méso-structurels. On notera des « espaces blancs » au sein des 'zones pavillonnaires'. Ils correspondent à l'introduction des 'espaces non construits' comme élément de caractérisation supplémentaire (cf. la construction d'indicateurs d'« insertion territoriale » p. 149).

Figure a.7.9. *Zone pavillonnaire ou zone d'habitat individuel*

Dans le cadre d'une application sur le suivi des *zones pavillonnaires*, il est plus rigoureux de parler de « zones d'habitat individuel » plutôt que de « zones pavillonnaires ». En effet la définition stricte d'un pavillon impose qu'il soit isolé (cf. p.209). Or comme on le constate sur les deux photographies de la figure a.7.9, une zone pavillonnaire, sous l'effet de la densification peut ne plus être une *zone pavillonnaire*, alors qu'elle demeure une *zone d'habitat individuel*.

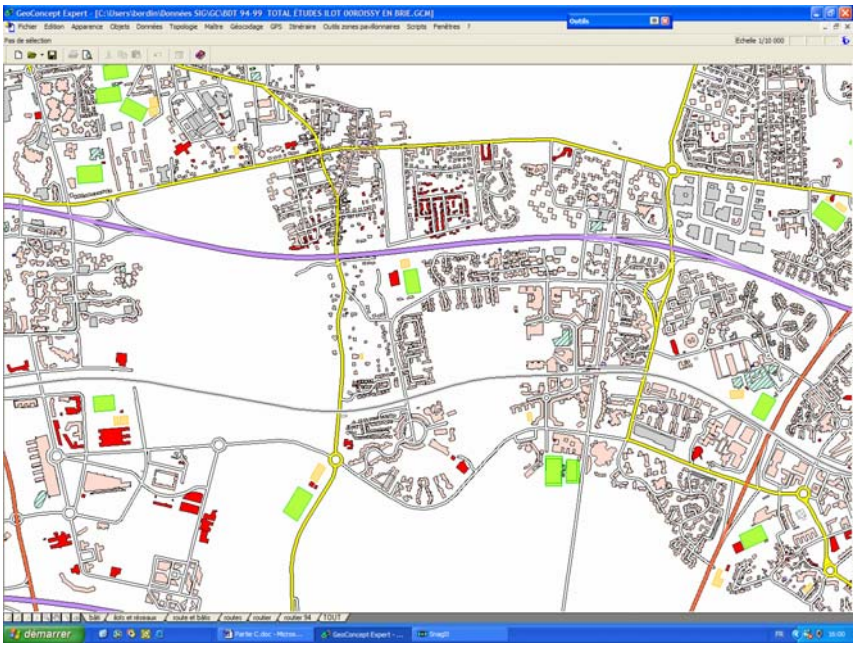


Figure a.7.10. *Identification automatique des créations de 'bâtiments'*

Le contexte spécifique de la *zone* étudiée localisée en ville nouvelle, nous a conduit à poser l'hypothèse que : toute création de 'bâtiment' se faisait sur un terrain non occupé précédemment par un autre 'bâtiment'. Sous cette hypothèse, nous avons alors effectué une analyse géométrique sur les 'bâtiments' présents dans la version de la BD TOPO® de 1999 qui ne se superposaient à aucun 'bâtiment' présent dans la version de 1993. Nous avons alors ainsi pu identifier un certain nombre de 'bâtiments' nouveaux, (identifiés en rouge dans l'extrait de carte de la figure a.7.10). Cette hypothèse ne serait évidemment pas valide en milieu urbain dense, où la création de nouveaux 'bâtiments' est essentiellement de type renouvellement urbain. En revanche dans le cas qui nous intéressait, elle s'est avérée tout à fait pertinente comme l'ont montré les résultats d'une étude comparative avec des données fournies par Arnaud Braun, issues d'une méthode d'extraction des évolutions plus élaborée et plus performante. Seulement 7% des 'bâtiments quelconques' nouveaux n'ont pas été identifiés par la méthode élémentaire mise en œuvre. Les figures qui suivent présentent deux exemples de création non identifiée par la méthode élémentaire proposée, que la méthode d'appariement ([Braun, 2004]) a permis de révéler. Elles montrent les limites de la méthode d'appariement utilisée qui repose sur une hypothèse de développement urbain très spécifique.

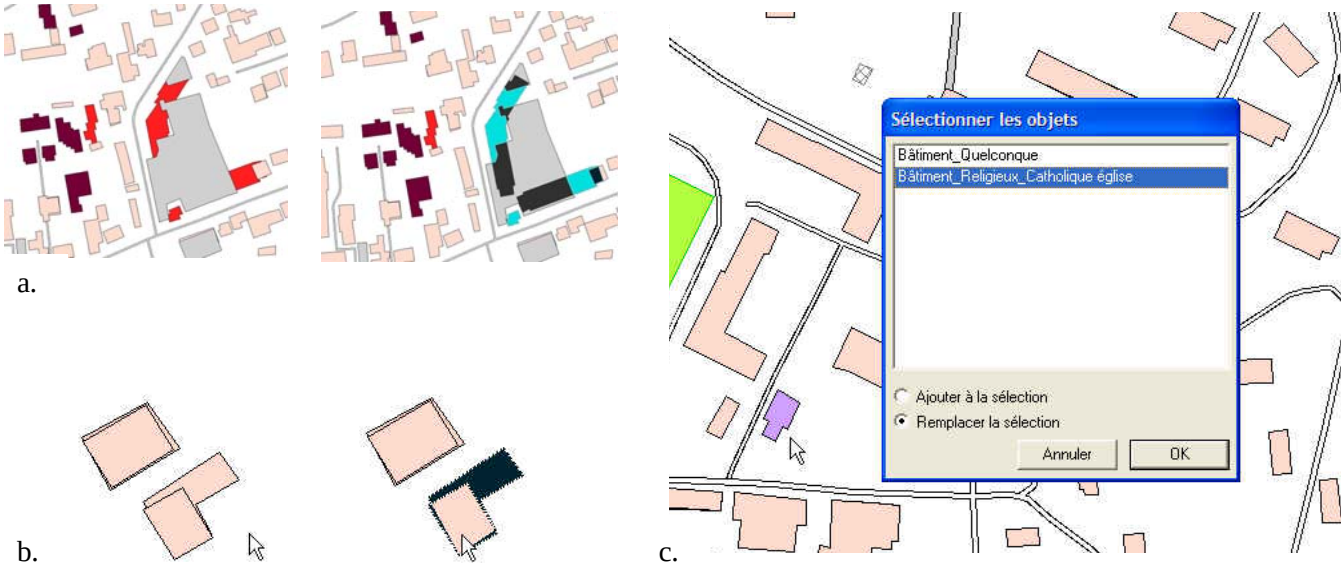


Figure a.7.11. *Méthode d'appariement pour l'identification des changements sur le bâti*

La méthode et les outils d'appariement mis en œuvre par Arnaud Braun ont permis d'identifier des créations de 'bâtiment' que la méthode par superposition ne permettait pas de déceler en raison de la superposition de bâtiments, cf. figure a.7.11(a). Elle a permis également de distinguer les modifications géométriques (b) (nous avons cependant montré dans le chapitre B.3 le risque d'une pré-interprétation des évolutions : s'agit t'il d'une modification ou d'une destruction/création, voire même d'une correction dans la base ?). Elle a permis aussi de saisir automatiquement des modifications sémantiques. Cependant, dans le cas de la BD TOPO® les informations sémantiques n'étant pas nombreuses, il y avait peu de possibilités de changement. On peut d'ailleurs parfois se demander si ceux identifiés correspondent à des modifications réelles ou bien s'il s'agit de mises à niveaux de la base (cf. A.3, p.71-72) comme dans le cas (c) : la méthode d'appariement à décelé une modification d'un 'bâtiment quelconque' en 'bâtiment religieux catholique église'. Ces défauts d'interprétation ne seraient alors pas imputables à la méthode d'extraction des changements, mais aux données elle-mêmes.

a.7.2. Eléments d'application pour l'observation de l'imperméabilisation de sols

Une partie importante de cette application repose sur la constitution des données d'échantillon. Il s'agit ici de bien répartir les surfaces imperméabilisées observées à partir d'une photographie aérienne en fonction des objets topographiques voisins, ceci afin d'obtenir les relations générales Ri les plus pertinentes.

Par exemple : dans le cas de la figure a.7.12, on observe en (a) cinq 'objets topographiques' de type 'bâtiment' situés dans une 'zone industrielle'. Il ne s'agissait pas de définir la 'zone imperméable' globale ('objet méso' pourtant proche de ce que l'on voulait observer) comme en (b), mais de tenter de distinguer celle associée à chaque 'bâtiment', ou du moins d'effectuer une répartition au plus juste des « surfaces imperméabilisées communes » entre les différents 'bâtiments' (c), pour contribuer à établir des relations Ri plus précises.

Figure a.7.12. *Stratégie d'acquisition des données d'échantillon pour l'élaboration des relations Ri servant au calcul des surfaces imperméables élémentaires (I).*

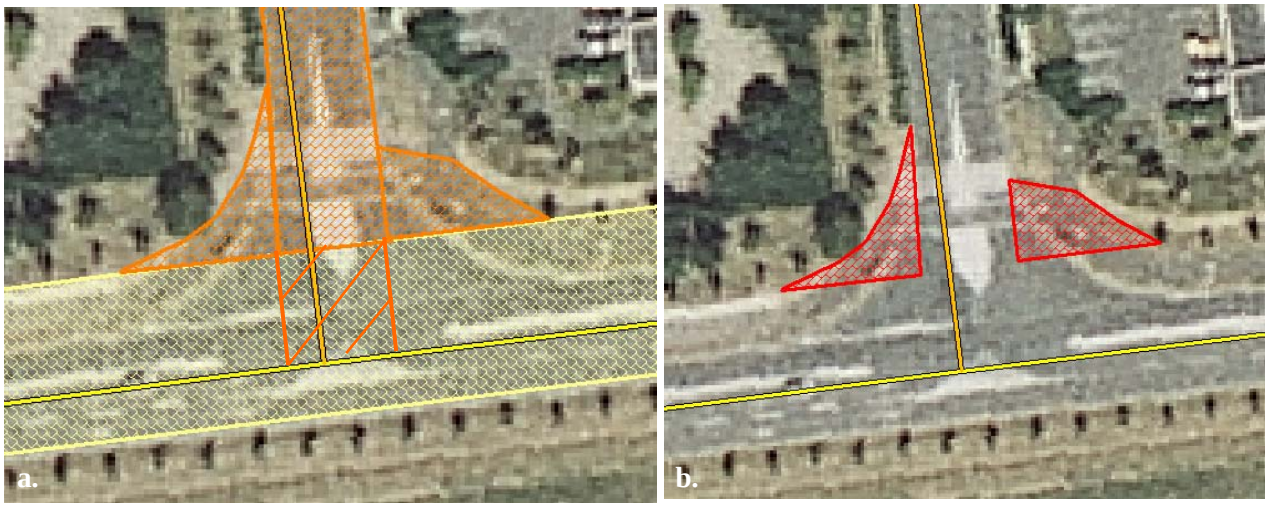


Figure a.7.13. *Stratégie d'acquisition des données d'échantillon pour l'élaboration des relations Ri servant au calcul des surfaces imperméables élémentaires (II).*

La 'surface imperméabilisée élémentaire' associée à un 'tronçon de route' est modélisée sous la forme d'un buffer symétrique autour de chaque 'tronçon'. L'objectif est de préciser la valeur du paramètre du buffer en fonction des caractéristiques des 'tronçons de route' telles que : la largeur de voie, le classement fonctionnel, etc. Or l'étude spécifique a montré que les carrefours introduisent deux types de biais (fig. a.7.13). D'une part, il existe une portion de 'surface imperméabilisée' qui est comptée deux fois : celle correspondant à l'intersection des buffers (a). D'autre part, la route ayant tendance à s'agrandir au niveau des carrefours (peut-être pour faciliter les virages), il existe une surface qui n'est prise en compte dans aucun des buffers, cette surface n'est donc pas comptée (b). En première approximation, nous avons fait l'hypothèse que les deux biais se compensaient. Des études plus détaillées pourraient être mises en œuvre pour analyser le cas spécifique des carrefours et tenter d'introduire des paramètres supplémentaires dans l'élaboration des Ri pour le 'réseau routier'.

a.7.3. Eléments d'application pour l'observation de l'ambiance urbaine

Les cartes qui suivent (fig. a.7.14 à a.7.16) font partie d'une série présentée dans [Basile, 2005]. Elles sont une illustration de la mise en application de la méthode d'observation spécifique élaborée pour observer l'ambiance urbaine. Elles montrent en particulier l'intérêt de la méthode spécifique pour mener des observations comparatives. Elles mettent ainsi en valeur le travail de formalisation technique réalisé pour construire des indicateurs à la fois reproductibles, transposables et robustes.

Figure a.7.14. *Cartes comparées sur les types de commerces présents* dans chacune des rues étudiées

Figure a.7.15. *Cartes comparées sur les horaires d'ouverture (mardi midi) dans les deux rues étudiées*

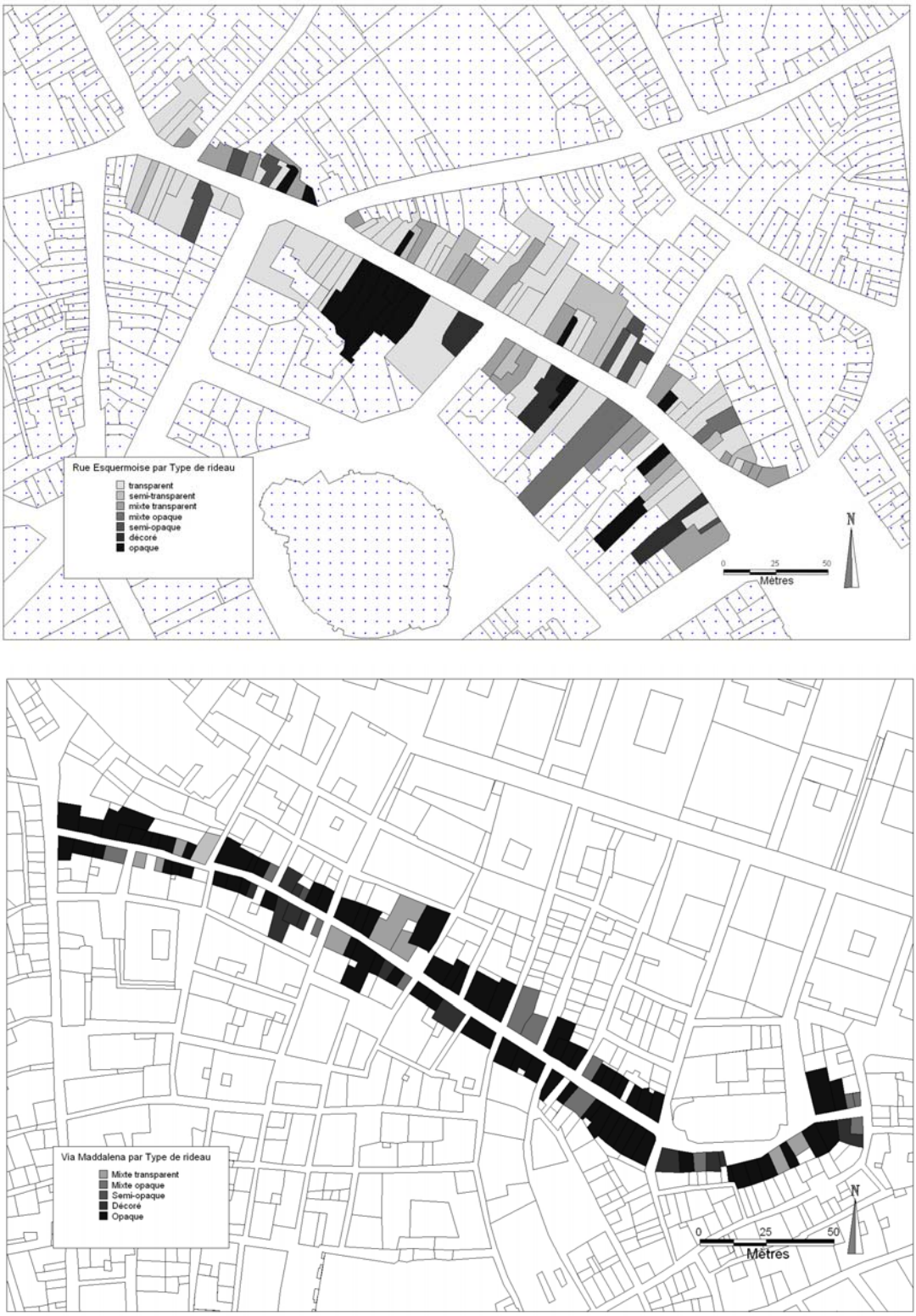


Figure a.7.16. *Cartes comparées sur la convivialité des deux rues étudiées en horaire de fermeture*

Annexe 8 : Etapes de mise en œuvre d'une méthode d'observation multi-niveaux

a.8.1. De la méthode générique aux applications

Pour aider au suivi de phénomènes géographiques, nous avons développé une méthode pour les observer dans les SIG : la *méthode d'observation multi-niveaux*. Cette méthode propose d'effectuer des observations statiques en établissant un lien entre plusieurs niveaux d'observation géographiques. Puis pour une observation cinématique, elle itère ces observations dans le temps en exploitant le lien chronologique entre les états. Elle s'appuie pour cela sur la méthode d'agrégation par emprise qui utilise un 'objet méso structurel' pour jouer le rôle d' « emporte-pièce ».

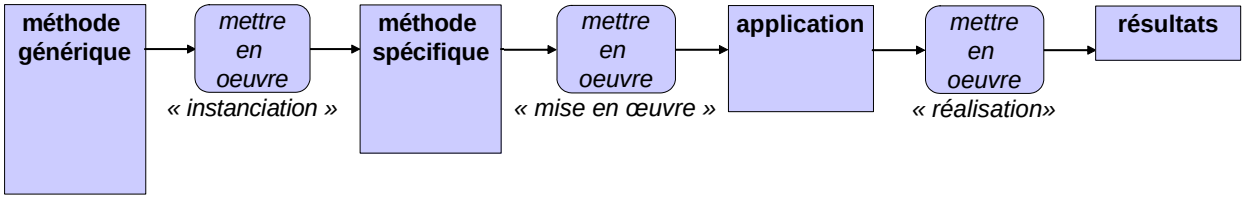


Figure a.8.1. *Mise en œuvre de la méthode générique d'observation multi-niveaux pour l'observation d'un phénomène spécifique sur une zone particulière.*

La méthode fonctionne selon le schéma de la figure a.8.1, avec trois stades de mise en œuvre : de la méthode générique jusqu'à l'obtention des valeurs observées (avant leur interprétation), en passant par la méthode spécifique. Le passage des uns aux autres s'effectue par mise en œuvre successive que nous différentions par des appellations distinctes :

– l'instanciation
– la mise en œuvre
– la réalisation

Enfin, chaque mise en œuvre se décline en différentes étapes.

a.8.2. Les étapes de mise en œuvre

a.8.2.1. L'instanciation

L'**instanciation** configure la méthode générique pour un phénomène particulier. Elle intègre le point de vue du thématicien sur le niveau et l'objet adéquats pour observer ce phénomène. Elle identifie parmi les données accessibles les informations micros et les caractéristiques utiles pour le suivi. Elle prépare les processus d'analyse des données et aboutit à l'élaboration d'une méthode spécifique dédiée au phénomène considéré (cf. figure a.8.2). A ce stade, la collaboration entre les acteurs est essentiel. En effet, pour tirer parti des données disponibles et produire des informations intéressantes pour le suivi, il est indispensable de mobiliser conjointement des compétences en données géographiques et des compétences thématiques spécifiques au phénomène.

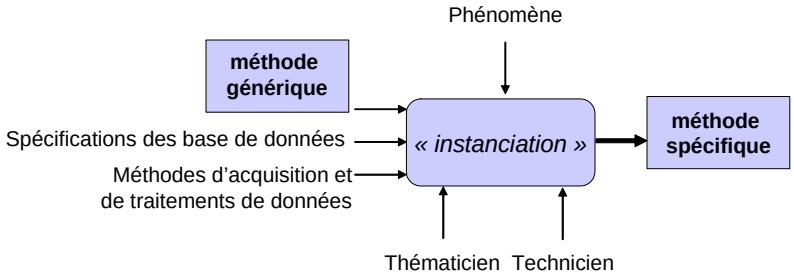


Figure a.8.2. *Instanciation de la méthode générique*

L'instanciation regroupe plusieurs étapes (fig. a.8.3), qui peuvent chacune se décliner en plusieurs questions (fig. a.8.4) :

⓪ Identification du niveau adéquat d'observation et de l'objet observé

L'identification du niveau et de l'objet adéquats d'observation est un travail préalable à la méthode. Il incombe au thématicien. Dans le paragraphe B.2.1 nous montrons comment ce travail préparatoire participe à la définition de l'objet (méso) géographique observé dans le SIG, sans toutefois se confondre avec elle.
Dans cette étape, il s'agit également de préciser les thèmes de caractérisation du phénomène (par exemple, pour les zones pavillonnaires, les thèmes peuvent être le bâti, la population, les accidents, etc.).

① Recensement des bases de données

Le recensement des bases de données sert à identifier les bases pouvant servir au suivi. Il prépare la sélection des données. Il importe ici de réunir des informations aussi bien sur leurs contenus (spécifications) que sur leur qualité, leur accessibilité ou leur modèle de gestion… Ces informations aideront par la suite à estimer le travail à faire sur les données. En particulier, les éléments collectés serviront pour spécifier les « systèmes de données » et exploiter les modèles de gestion temporelle.

② Etude des données

Le recueil d'informations sur les données disponibles[7] permet de :

- spécifier l'emprise d'agrégation et définir l'objet méso structurel utilisé.
- identifier les données exploitées au niveau micro.

Cette étape nécessite à la fois des compétences thématiques et des compétences en données géographiques. Elle réalise des allers-retours entre les deux pour configurer les données en fonction de l'application thématique. Ceci est particulièrement vrai lorsque les données n'offrent pas d'objets méso structurels naturels (voir l'application à l'ambiance urbaine), mais reste valable, de façon générale pour sélectionner les données micro (par exemple, quelles données intégrer pour exprimer la convivialité des commerces ?).

③ Sélection des indicateurs

Dans cette étape, il s'agit de préciser les indicateurs utilisés pour caractériser le phénomène. Cette étape prolonge la précédente. Elle affine le choix des caractéristiques et détermine la façon de les obtenir. Les indicateurs pourront être classés par type d'analyse mise en jeu (pour préparer les développements de la méthode spécifique) et les ressources pour les calculer (logiciels, expertises) pourront être recensées.

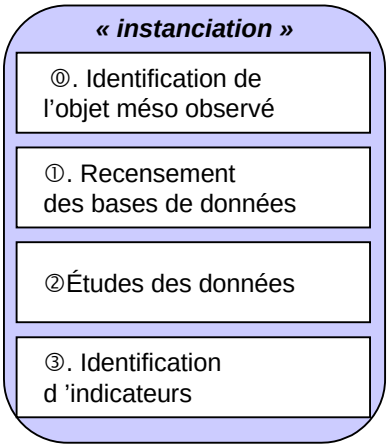

« instanciation »
⓪. Identification de l'objet méso observé
①. Recensement des bases de données
②Études des données
③. Identification d 'indicateurs

Figure a.8.3. *Les étapes de l'instanciation*

Exemple de questions traitées dans les étapes de l'instanciation :

① Quelles sont les données qui pourraient fournir une représentation de l'objet méso (ou d'éléments de partition adéquate)?
Quelles bases de données pourraient contenir des informations intéressantes sur le phénomène à un niveau micro ?
Quelles sont leurs spécifications ?
Sous quelles formes (quelle modélisation) les informations sont elles représentées (localisées, localisables, par nature) ?
Avec quelle précision ? Quelles données support ? Quelles qualités ?
Quel modèle de gestion temporelle est utilisé pour chaque base? Dates des mises à jour ?

② Combien de systèmes de données doit-on construire ? Quels types d'objets méso structurels pour chacun ?
Quelles données micro ? Quels indicateurs permettront-elles de construire ?

③ Quels indicateurs ? Sont-ils statistiques, géométriques, topologiques ? Où trouver des moyens pour les calculer ?…

Figure a.8.4. *Les questions traitées lors de l'instanciation*

7. *Disponible* est à prendre ici dans le sens large. Il regroupe l'ensemble des données que l'utilisateur peut mobiliser pour son observation.

a.8.2.2. La mise en œuvre

La **mise en œuvre** fixe les différents éléments de la méthode. Elle s'appuie sur la méthode spécifique et intègre les données et moyens disponibles pour construire l'application. Elle détaille comment construire l'emprise d'agrégation lorsqu'elle n'est pas « donnée ». Elle spécifie le travail d'adaptation et d'intégration des données micro. Enfin, elle indique comment obtenir chaque indicateur. Les différents choix tiennent compte des moyens de l'utilisateur final, pour assurer l'adéquation de l'application résultante (fig. a.8.5).

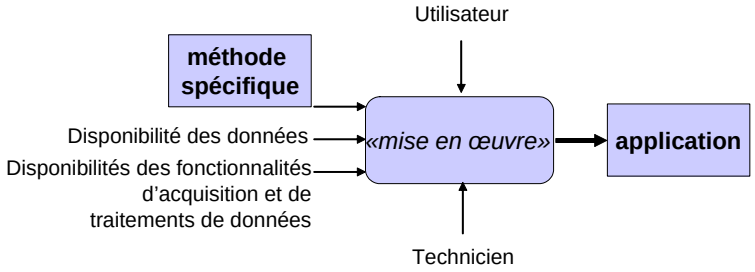


Figure a.8.5. *Mise en œuvre d'une méthode spécifique*

La *mise en œuvre* comprend trois étapes (fig. a.8.6) qui peuvent se décliner sous forme d'une série de questions (fig. a.8.7) :

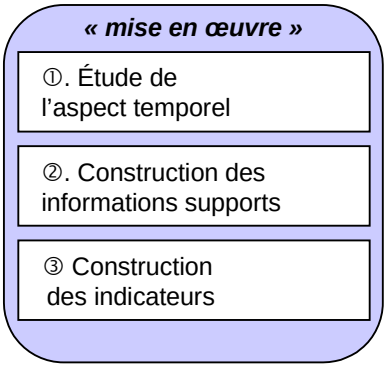


Figure a.8.6. *Les étapes de la mise en oeuvre*

① Etude de l'aspect temporel

L'étude des données lors de l'*instanciation* fournit les éléments pour l'observation statique du phénomène. Il s'agit ici d'intégrer la dimension temporelle. Cette étape clarifie le processus d'observation en fonction des modèles de gestion et des dates des mises à jour, en fonction des systèmes de données et d'éventuelles fonctions d'interpolation des états.

② Construction des informations supports : objet méso structurel et informations micro

La *construction des informations support* est l'étape qui prépare les données pour l'application de la méthode d'observation multi-niveaux.

– Elle spécifie le travail à effectuer pour construire l'emprise d'agrégation. Le processus proposé peut être élémentaire (ex. : simple requête sur la composante géométrique de l''objet méso géographique') ou renvoyer à des développements complémentaires. (Par exemple, dans le cas des zones pavillonnaires nous avons testé une méthode pour construire des objets méso géographiques spécifiques géométriquement plus précis que les autres alternatives : *zone pavillonnaire du MOS* et *îlots routiers*).

– Cette étape sert également à configurer les données micro. Elle opère sur les données pour disposer des informations nécessaires au calcul des indicateurs dans le SIG. Il peut s'agir de :

- modifier le modèle de données existantes, par exemple : transformation d'informations localisées en informations par nature (voir le travail sur les zones pavillonnaire du MOS p. 108)
- acquérir des données nouvelles (par géocodage de fichiers non géographiques, par enquêtes terrain ou par numérisation, etc.)
- effectuer des analyses préalables pour disposer d'attributs micro supplémentaires et permettre ainsi de calculer les indicateurs méso correspondants. Par exemple, le calcul préliminaire d'une caractéristique de forme pour le bâti (ex. : la concavité) permet ensuite

Exemple de questions traitées dans les étapes de mise en œuvre :

① Quels sont les modèles de gestion temporelle associés à chaque base ? Peut-on mettre en place une observation à plusieurs niveaux emboîtés ? Combien de systèmes de données doit-on définir ? Quelles sont les dates de mise à jour ? Sont-elles compatibles ? Si non, utilise t'on une fonction d'interpolation et laquelle ?

② Existe t'il une donnée représentant l'objet méso géographique ? Sa composante peut-elle servir pour l'agrégation (est-elle compatible avec les systèmes de données principaux) ? Si non, quelle stratégie adopter pour définir l'emprise d'agrégation (cf. B.2.4.3) ?

Figure a.8.7. (première partie) *Les questions traitées lors de la mise en œuvre*

de créer un indicateur sur la morphologie du bâti des zones pavillonnaires.

③ Construction des indicateurs

Il s'agit ici de construire la famille f_k des indicateurs de suivi, en tenant compte des ressources et moyens prévus pour l'utilisateur final. Dans cette étape le processus de calcul de chaque indicateur est défini. Les fonctionnalités requises sont spécifiées et les outils nécessaires sont éventuellement développés. En particulier, les analyses préalables pour introduire des attributs micro supplémentaires sont mises en œuvre et les outils pour calculer leur valeur sont mis au point.

② (suite) Les données micro donnent-elles un accès direct aux informations nécessaires dans le SIG ? Faut-il adapter leur modélisation ? Manque t'il des informations ? Où les trouver ? Faut-il les produire ?

③ Quelle est la liste des indicateurs ? Les fonctionnalités nécessaires à leur calcul sont-elles disponibles ? Si non où les trouver ? Doit-on les développer ?…

Figure a.8.7. (seconde partie) *Les questions traitées lors de la mise en œuvre*

a.8.2.3. La réalisation

La **réalisation** effectue les opérations conçues dans la méthode spécifique sur des extraits (ou jeux) de données correspondant à la zone étudiée. Elle intègre les données micro dans le SIG et les configure pour l'observation. Elle applique la méthode d'agrégation par emprise en utilisant l'emprise préalablement définie et réalise les calculs de caractéristiques. Pour aider au suivi, elle peut expliciter les variations de chaque caractéristique pour tous les objets observés, préparant ainsi l'interprétation des indicateurs et l'analyse du phénomène.

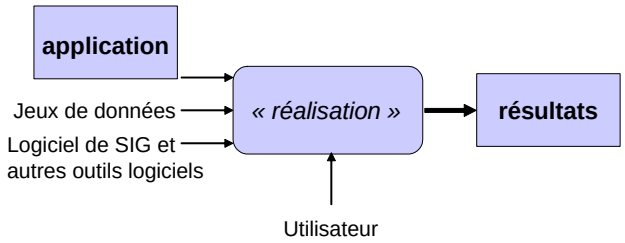


Figure a.8.8. *Réalisation d'une observation multi-niveaux*

Dans cette étape l'utilisateur s'approprie la méthode d'observation (fig. a.8.8). Après l'avoir expérimentée, il pourra demander des développements complémentaires (ex. : introduction de nouveaux indicateurs, de nouveaux thèmes, amélioration de la précision sur les indicateurs, de stratégie temporelle plus sophistiquée, etc.). En fonction du besoin, il s'agira de revenir à l'une des étapes antérieures de mise en œuvre. Par exemple, le besoin d'une plus grande précision sur les indicateurs peut impliquer de revenir à l'étape de *construction des indicateurs* (dans la « mise en œuvre ») pour effectuer des développements supplémentaires ou bien impliquer de revoir le processus dès l'étape de *recensement des données* (dans l'« instanciation ») pour introduire des données micro plus précises. La *réalisation* comprend cinq étapes (fig. a.8.9) :

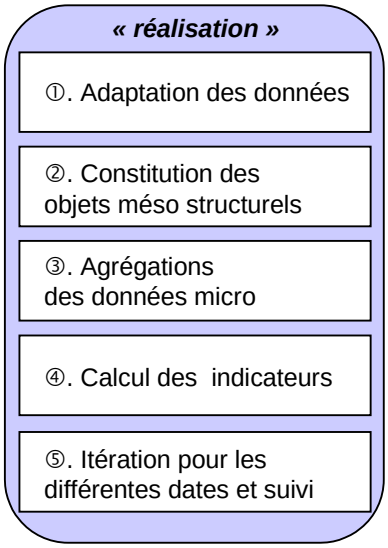

« réalisation »
①. Adaptation des données
②. Constitution des objets méso structurels
③. Agrégations des données micro
④. Calcul des indicateurs
⑤. Itération pour les différentes dates et suivi

Figure a.8.9. *Les étapes de la réalisation*

① Adaptation des données

Au cours de l'*adaptation des données*, l'utilisateur intègre les différents jeux de données. Il effectue les pré-traitements qui lui permettront de disposer dans le SIG des informations micro nécessaires pour l'observation multi-niveaux (ex. : introduction des dates de version dans les bases possédant un modèle par snapshots sans estampillage, modification de structure des données pour introduire de nouveaux attributs, enrichissement des données par géocodage ou autre numérisation, etc.).

② Constitution des objets méso structurels

Dans cette étape, l'utilisateur constitue les emprises d'agrégation. Les opérations à réaliser dépendent de la stratégie d'agrégation et surtout de la disponibilité de données pour représenter l'objet méso-structurel (cf. B.2.5).

③ Agrégation des objets micro par emprise

Dans cette étape, l'utilisateur applique la méthode d'agrégation par emprise en s'appuyant sur le SIG : Il exploite les outils de sélection géométrique (fonctionnalité standard des SIG), pour construire à partir des objets méso structurels, les agrégations de données micro.

④ Calcul des indicateurs

Les méthodes et outils de calculs des indicateurs ayant été mis au point, au stade précédent, l'utilisateur n'a plus ici qu'à les appliquer aux agrégations de données micro pour calculer les valeurs des indicateurs.

⑤ Itération sur les différentes dates et suivi

Les étapes précédentes permettent pour une date fixée de calculer les caractéristiques des 'objets méso'. Pour effectuer un suivi, l'utilisateur doit itérer les étapes ③ et ④ pour les autres dates observées. Au final, il obtient des séries temporelles d'indicateurs sur les différents 'objets méso' observés, séries qu'il peut ensuite analyser, interpréter… et étendre au gré des mises à jour.